

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.594 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-39
Perfect score: 107
Sequence: 1 SHICRRPKYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA2001.DAT.*
23: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA2002.DAT.*
24: /SIDSL/gcgdata/geneeq/geneeqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	107	100.0	20 18 AAM42922	Immunogenic Hepat
2	107	100.0	21 22 AAB6439	Synthetic HAV P2A
3	107	100.0	366 6 AAP50230	Sequence of Hepat
4	107	100.0	854 6 AAP50287	Sequence encoded b
5	107	100.0	993 6 AAP50116	Sequence of Hepat
6	107	100.0	993 6 AAP50231	Sequence encoded b
7	107	100.0	1077 20 AAW95559	A partial hepatitis
8	107	100.0	1091 14 AAR37426	Translated from 5'
9	107	100.0	2227 11 AAR05697	Attenuated hepatitis

10	107	100.0	2227	18	AAM34074	Hepatitis A virus
11	107	100.0	2227	21	AAB18607	Amino acid sequenc
12	107	100.0	2227	21	AAB18608	Amino acid sequenc
13	107	100.0	2227	21	AAB18609	Amino acid sequenc
14	107	100.0	2227	23	ABG31727	Wild-type Hepatit
15	107	100.0	2227	23	ABG31728	Hepatitis A virus
16	107	100.0	2227	23	ABG31729	Hepatitis A virus
17	107	100.0	2227	24	AAE18989	Attenuated Hepatit
18	107	100.0	2227	24	ABU08639	Hepatitis A virus
19	107	100.0	2227	24	ABU08640	Wild type human he
20	107	100.0	2227	24	ABU08641	Attenuated (pass3
21	91	85.0	2227	7	AP60066	Sequence of viral
22	84.5	79.0	839	12	AAE15629	Capid region of c
23	67	62.6	20	18	AAW42923	Immunogenic Hepat
24	67	62.6	21	22	AAW42923	Synthetic HAV P2A
25	45.5	42.5	178	22	AAU03662	Group B Streptococ
26	45.5	42.5	330	23	ABP30218	Streptococcus poly
27	45.5	42.5	341	23	ABP25499	Human polypeptide
28	45	42.1	1273	23	ABP62888	Human polypeptide
29	44	41.1	71	21	AAW07411	Arbidopsin thalia
30	44	41.1	88	21	AAW61775	Arbidopsin thalia
31	44	41.1	211	22	AAW81246	Human AFP protein
32	44	41.1	346	12	AAE13068	12D3 antigen seque
33	43	40.2	45	22	AAE13255	A tumour necrosis
34	43	40.2	45	22	AAE13255	Human tumour necro
35	43	40.2	52	22	AAE13255	Human tumour necro
36	43	40.2	52	22	AAE13255	A tumour necrosis
37	43	40.2	287	21	AAW06837	Human tumour necro
38	43	40.2	312	21	AAW06837	Arbidopsin thalia
39	43	40.2	320	21	AAW06836	Arbidopsin thalia
40	43	40.2	332	23	ABE91335	Arbidopsin thalia
41	43	40.2	372	22	ABE71552	Herbicide-inducible
42	43	40.2	391	13	AAE70479	Drosophila melanog
43	43	40.2	391	13	AAE70479	Sequence of the ma
44	43	40.2	391	19	AAW25305	HRSV major capsid
45	43	40.2	413	22	AAE84220	HRSV major nucleoc

ALIGNMENTS

RESULT 1
ID AAM42922 standard; peptide; 20 AA.
XX
AC AAM42922:
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1315.
XX
DE Immunogenic peptide; immunogenic epitope; P2A protein;
XX
KW Immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE.
XX
PI Fields HA, Khudyakov YE.
XX
DR WPI: 1997-535631/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT Immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-811. The present peptide
CC is derived from amino acids 792-811, and has a reactivity of 54.28 with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
SQ Sequence 20 AA:
Query Match 100.0%; Score 107; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SHIECRKPYKELRLEVGKOR 20
1 SHIECRKPYKELRLEVGKOR 20
Db
RESULT 2
AAB69439
ID AAB69439 standard; Peptide; 21 AA.
XX
AC AAB69439;
XX
DF 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 39.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
PY 15-JUL-1999; 99US-0144412.
XX
FR (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 93; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
SQ Sequence 21 AA:
Query Match 100.0%; Score 107; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SHIECRKPYKELRLEVGKOR 20
1 SHIECRKPYKELRLEVGKOR 20
Db
RESULT 3
AAP50230
ID AAP50230 standard; Protein; 366 AA.
XX
AC AAP50230;
XX
DF 28-NOV-1991 (first entry)
XX
DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
OS Hepatitis A virus.
XX
FN EPI38704-A.
XX
PD 24-APR-1985.
XX
PF 09-OCT-1984; 84EP-0402025.
XX
PR 02-MAR-1984; 84US-0585942.
PR 14-OCT-1983; 83US-0541836.
XX
PA (MERI) MERCK & CO INC.
XX
PI Hughes JV, Scolnick EM, Tomassini JE;
XX
DR WPI; 1985-100818/17.
XX
DR N-PSDB; AAN50274.
XX
PT New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus
XX
PS Claim 21; Page 46-48; 49pp; English.
XX
CC VP1 is isolated by solubilisation of the intact virus in an aq.
CC anionic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ Sequence 366 AA:
Query Match 100.0%; Score 107; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SHIECRKPYKELRLEVGKOR 20
1 SHIECRKPYKELRLEVGKOR 320
Db
RESULT 4
AAP50287
ID AAP50287 standard; Protein; 854 AA.
XX
AC AAP50287;
XX


```

DT 25-MAR-2003 (updated)
DT 30-NOV-1991 (first entry)
XX
XX
DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the
DE genome 5' terminus to the end of the area corresponding to the
DE capsid protein region of poliovirus RNA.
XX
XX
KM Hepatitis A virus assay; antigen; antibody.
XX
XX
OS Hepatitis A virus.
XX
XX
PN W06501517-A.
XX
XX
PD 11-APR-1985.
XX
XX
PF 27-SEP-1984; 84MO-US01552.
XX
XX
PR 30-SEP-1983; 83US-0537911.
XX
XX
(MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX
PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
PI Racanello VR.
XX
XX
DR WPI: 1985-098846/16.
DR N-PSDB; AAN50330.
XX
XX
PT New hepatitis A virus cDNA - useful in assays for the virus and
PT for prodn. of the viral antigen and antibodies to it
XX
XX
PS Example: Fig 7; 60pp; English.
XX
XX
CC The inventors claim HAV cDNA and a method for producing it, whereby
CC large amts. can be obtd. economically. The cDNA is useful in the
CC assay for detection of HAV quickly and easily and with high
CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX
SQ Sequence 854 AA;
XX
XX
Query Match 100.0%; Score 107; DB 6; Length 854;
Best Local Similarity 100.0%; Pred. No. 5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SHIECRKPKYKELREVGKOR 20
DB 792 SHIECRKPKYKELREVGKOR 811
XX
XX
RESULT 5
AAP50116
ID AAP50116 standard; Protein: 993 AA.
XX
XX
AC AAP50116;
XX
XX
DT 25-MAR-2003 (updated)
DT 30-SEP-1991 (first entry)
XX
XX
DE Sequence of Hepatitis A virus (HAV) immunogenic peptides
DE VP-1, VP-2, VP-3 and VP-4.
XX
XX
KM Antigenic protein; immunogen; vaccine.
XX
XX
OS Hepatitis A virus (strain CR326).
XX
XX
PN EPI54587-A.
XX
XX
PD 11-SEP-1985.
XX
XX
PF 27-FEB-1985; 85SEP-0400369.
XX
XX
PR 02-MAR-1984; 84US-0585818.
XX
XX

```

```

XX
XX
PA (MERI ) MERCK & CO INC.
XX
XX
PI Linemeyer DL, Menke JG, Rueben RG, Mltra SW;
XX
XX
DR WPI: 1985-224964/37.
XX
XX
DR N-PSDB; AAN50139.
XX
XX
PT New nucleotide sequences coding for hepatitis A virus antigens -
PT useful for eliciting normal immune response and in vaccines for
PT protecting against the virus
XX
XX
PS Example: Page 11-17; 32pp; English.
XX
XX
CC Within the sequence in AAN50139 is encoded the information necessary
CC to make the antigenic proteins of HAV. The sequences encoding for
CC the structural proteins begin at base 403. The key sub-unit
CC sequences within VP-1, designated Sequences I, II, III, IV, and V,
CC start, respectively at 1882, 1963, 1999, 2146, 2347. Other
CC nucleotide sequences which are valuable as encoding antigenic
CC proteins are the sequences from base 1749 to base 2722; from base
CC 1487 to base 2980 and from base 1644 to base 2722. The sequence from
CC base 1749 to base 2722 is esp. valuable as a vector for producing
CC antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
CC translation of a stop codon.
XX
XX
SQ Sequence 993 AA;
XX
XX
Query Match 100.0%; Score 107; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SHIECRKPKYKELREVGKOR 20
DB 928 SHIECRKPKYKELREVGKOR 947
XX
XX
RESULT 6
AAP50231
ID AAP50231 standard; Protein: 993 AA.
XX
XX
AC AAP50231;
XX
XX
DT 28-NOV-1991 (first entry)
XX
XX
DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
XX
XX
KM Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KM diagnostic assay.
XX
XX
OS Hepatitis A virus.
XX
XX
FH Key Location/Qualifiers
FH FT 628..993
FH FT /note="claimed; X denotes translated stop codons
FH FT and unspecified triplets"
XX
XX
PN EPI38704-A.
XX
XX
PD 24-APR-1985.
XX
XX
PF 09-OCT-1984; 84EP-0402025.
XX
XX
PR 02-MAR-1984; 84US-0585942.
XX
XX
PR 14-OCT-1983; 83US-0541836.
XX
XX
PA (MERI ) MERCK & CO INC.
XX
XX
PI Hughes JV, Scolnick EM, Tomassini JE;
XX
XX
DR WPI: 1985-100818/17.
XX
XX

```

DR N-PSDB; AAN50274.
 XX
 PT New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus
 XX
 PS Disclosure; Page 17-23; 49pp; English.
 CC
 CC vpi is isolated by solubilisation of the intact virus in an aq.
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.
 CC and the protein of molecular wt. 33000 daltons is sepd.
 XX
 SO Sequence 993 AA;
 Query Match 100.0%; Score 107; DB 6; Length 993;
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIECRKPKYKELRLEVGKOR 20
 |||||
 928 SHIECRKPKYKELRLEVGKOR 947
 RESULT 7
 AAM95559 standard; Protein; 1077 AA.
 AC AAM95559;
 XX
 DT 28-APR-1999 (first entry)
 XX
 DE A partial hepatitis A virus (HAV) protein.
 XX
 KW Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX
 OS Hepatitis A virus.
 XX
 PN US5849562-A.
 XX
 PD 15-DEC-1998.
 XX
 PF 06-JUN-1995; 95US-0468926.
 XX
 PR 06-NOV-1991; 91US-0788262.
 PR 30-SEP-1983; 83US-0537911.
 PR 27-SEP-1984; 84US-0654942.
 PR 06-OCT-1988; 88US-0256135.
 PR 06-JUN-1995; 95US-0468926.
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Emerson SU, Purcell RH;
 DR WPI; 1999-094412/08.
 DR N-PSDB; AAX01006.
 XX
 PT Chimeric hepatitis A virus strains - with P2 region from
 PT cell-culture-adapted strain in wild-type genome
 XX
 PS Disclosure; Fig 7A-L; 36pp; English.
 CC
 CC The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting
 CC of a wild-type HAV genome in which the P2 region is replaced by the
 CC P2 region from a cell-culture-adapted HAV strain. The construct is
 CC used to demonstrate that mutations in the P2 region of a
 CC cell-culture-adapted HAV strain are sufficient for establishment of
 CC infection and accelerated growth in cell culture.
 CC
 SO Sequence 1077 AA;
 Query Match 100.0%; Score 107; DB 20; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 6.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIECRKPKYKELRLEVGKOR 20
 |||||
 1015 SHIECRKPKYKELRLEVGKOR 1034
 Db
 RESULT 8
 AAR32426
 ID AAR32426 standard; Protein; 1091 AA.
 XX
 AC AAR32426;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-DEC-2001 (updated)
 DT 10-JUN-1993 (first entry)
 XX
 DE Translated from 5' region of Hepatitis A virus genomic clone.
 XX
 KW HAV HM-175; chronic liver disease; picornavirus.
 XX
 OS Hepatitis A Virus.
 XX
 FH Key Location/Qualifiers
 FT Region 238..1091
 FT /label= ORF
 FT /note= "second putative initiation codon at
 FT position 240"
 FT Region 1..711
 FT /note= "X's correspond to nonsense codons,
 FT 1.e. this region is not an ORF"
 XX
 PN USN7788262-N.
 XX
 PD 15-DEC-1992.
 XX
 PF 30-SEP-1983; 83US-0536911.
 XX
 PR 27-SEP-1984; 84US-0654942.
 PR 06-OCT-1988; 88US-0256135.
 PR 30-SEP-1983; 83US-0536911.
 PR 06-NOV-1991; 91US-0788262.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PI Ticehurst JR, Baltimore D, Felinstone SM, Purcell RH, Racanietello VR;
 PI Baroudy BM, Emerson SU;
 DR WPI; 1993-067429/08.
 DR N-PSDB; AAQ36934.
 XX
 PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
 PT of antigen and antibodies
 XX
 PS Disclosure; Fig 7; 65pp; English.
 CC
 CC HAV virion RNA was extracted from the livers of marmosets which had
 CC been inoculated with HAV (the HAV had previously been passaged twice
 CC in marmosets). The RNA was used to prepare ds cDNA clones by
 CC standard methods. Clones contg. inserts which hybridised to RNA from
 CC HAV-infected African Green Monkey kidney cells were selected for
 CC further analysis. A 7.4kb restriction map (about 99% of the HAV
 CC genome) was constructed from 5 overlapping inserts. The sequence of
 CC the first 3.3kb (approx.) from the 5'-terminus was determined. An
 CC amino acid sequence was deduced from the entire clone and an open
 CC reading frame was identified starting at position 238. A comparison
 CC of the predicted HAV amino acid sequences with the known capsid
 CC protein sequences of other picornaviruses (poliovirus, foot and
 CC mouth disease virus and encephalomyelitis virus) revealed areas of
 CC local homology.
 CC (Note: Revised entry submitted to correct the patent number format of
 CC US Government-owned NTIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent

CC web site at www.derwent.com/dwpi/updates/ntis-us.html.)
 CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 1091 AA;

Query Match 100.0%; Score 107; DB 14; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 6.5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVKGKOR 20
 ||||||||||||||||||

DB 1029 SHIECRKPYKELRLVKGKOR 1048

RESULT 9

AAW05697
 ID AAW05697 standard; protein; 2227 AA.

XX AAW05697;

25-MAR-2003 (updated)
 15-AUG-1990 (first entry)

XX Attenuated hepatitis A virus.

XX Hepatitis A virus; vaccine; attenuated.

OS Hepatitis A virus, strain HM-175.

PH Key Location/Qualifiers

FT 1..23 /label=VP4 - 1A

FT 24..245 /label=VP2 - 1B

FT 246..491 /label=VP3 - 1C

FT 492..791 /label=VP1 - 1D

FT 792..980 /label=2A

FT 981..1087 /label=2B

FT 1088..1422 /label=2C

FT 1423..1496 /label=3A

FT 1497..1519 /label=3B - VPg

FT 1520..1738 /label=3C

FT 1739..2227 /label=3D

XX Region

XX US4894228-A.

XX 16-JAN-1990.

XX 12-JUL-1988; 88US-0217824.

XX 12-JUL-1988; 88US-0217824.

XX 12-JUL-1988; 88US-0652967.

XX (USSH) US DEPT HEALTH & HUMAN SERVICE.

XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM.

XX Daemer RJ, Gust ID;

XX WPI; 1990-075557/10.

XX N-PSDB; AAQ03512.

XX Vaccine against hepatitis A virus infection - comprises novel

XX attenuated hepatitis A virus strain.

PS Claim 1; Fig 1; 18pp; English.

XX

CC The attenuated HAV is useful for inducing protective immunity against

CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by

CC several nucleotide changes distributed throughout the genome, is

CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is

CC suitable for use as an HAV vaccine. It is noted that not all the changes

CC are necessary for attenuation and use as a vaccine.

CC (Updated on 25-MAR-2003 to correct PA field.)

CC (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 11; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVKGKOR 20
 ||||||||||||||||||

DB 792 SHIECRKPYKELRLVKGKOR 811

RESULT 10

AAM34074
 ID AAM34074 standard; protein; 2227 AA.

XX AAM34074;

27-APR-1998 (first entry)

XX Hepatitis A virus HM-175 protein sequence.

XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;

XX infection; vaccine.

XX Hepatitis A virus HM-175.

PH Key Location/Qualifiers

FT 1..23 /label=VP4

FT 24..245 /label=VP2

FT 246..491 /label=VP3

FT 492..791 /label=VP1

FT 792..980 /label=2A

FT 981..1087 /label=2B

FT 1088..1422 /label=2C

FT 1423..1496 /label=3A

FT 1497..1519 /label=3B

FT 1520..1738 /label=3C

FT 1739..2227 /label=3D

XX Protein

XX WO9740166-A2.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US06506.

XX 19-APR-1996; 96US-0015642.

XX (USSH) US SEC DEPT HEALTH.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Emerson SU, Purcell RH, Raychaudhuri G;

XX WPI: 1997-535850/49.
DR N-PSDB: AAT93023.
XX
PT Human attenuated HAV genome containing simian HAV 2C gene - useful
PT as vaccines against HAV infection
XX
PS Disclosure: Fig 13A-D; 66pp; English.
XX
CC This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (I) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKOR 20
DB 792 SHIECRPKYKELRLEVGKOR 811
RESULT 11
AAB18607
ID AAB18607 standard; Protein: 2227 AA.
XX
AC AAB18607;
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX
HA; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93MO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI: 2000-586464/55.
DR N-PSDB: AAA75476.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
PS Disclosure: Fig 6A-K; 72pp; English.

XX
CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKOR 20
DB 792 SHIECRPKYKELRLEVGKOR 811
RESULT 12
AAB18608
ID AAB18608 standard; Protein: 2227 AA.
XX
AC AAB18608;
XX
DT 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX
HA; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX
P-35 virus.
XX
OS Hepatitis A virus.
XX
PN US6113912-A.
XX
PD 05-SEP-2000.
XX
PF 07-JUN-1995; 95US-0475886.
XX
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93MO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR WPI: 2000-586464/55.
DR N-PSDB: AAA75477.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
PS Disclosure: Columns 67-78; 72pp; English.
XX
CC The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKOR 20
DB 792 SHIECRPKYKELRLEVGKOR 811

Db 792 SHIECRKPKYKELRLEVGNOR 811

RESULT 13

AA18609
ID AAB18609 standard; Protein: 2227 AA.

AC AAB18609;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

KM HAV 4380.

OS Hepatitis A virus.

PN US6113912-A.

PF 05-SEP-2000.

PR 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

DR WPI; 2000-586464/55.

DR N-PSDB; AAA75478.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type

PS Disclosure; Columns 93-104; 72pp; English.

CC The present sequence is derived from a live attenuated hepatitis A

CC virus (HAV) of the invention, designated HAV 4380. The sequence is

CC produced by modifying wild type HAV strain HM-174. The HAV of the

CC invention are adapted to growth in the human fibroblast-like cell

CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain

CC appropriate attenuation. It is useful as a live vaccine for prophylaxis

CC of hepatitis A in humans and other primates.

CC Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGNOR 20

DB 792 SHIECRKPKYKELRLEVGNOR 811

RESULT 14

ABG31727

ID ABG31727 standard; Protein: 2227 AA.

AC ABG31727;

DT 29-NOV-2002 (first entry)

DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

KM Hepatitis A virus strain HM-175.

XX US6423318-B1.

PN 23-JUL-2002.

PD 31-AUG-2000; 2000US-0653499.

PF 07-JUN-1995; 95US-0475886.

PR 17-SEP-1993; 93US-0397232.

PR 17-SEP-1993; 93WO-US08610.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

DR WPI; 2002-680946/73.

DR N-PSDB; ABS52787.

PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection

PS Disclosure; Fig 6; 71pp; English.

CC The invention relates to a polynucleotide which encodes a hepatitis A

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

CC line). The polynucleotide is useful for preparing a vaccine against

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

CC strain HM-175 polypeptide.

CC Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGNOR 20

DB 792 SHIECRKPKYKELRLEVGNOR 811

RESULT 15

ABG31728

ID ABG31728 standard; Protein: 2227 AA.

AC ABG31728;

DT 29-NOV-2002 (first entry)

DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.

KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

KM virucide; mutant; PHAV/7; mutlein.

OS Hepatitis A virus strain HM-175.

OS Synthetic.

Key Location/Qualifiers

FT MISC-difference 963

FT MISC-difference 764

FT MISC-difference 821

FT MISC-difference 1052

FT MISC-difference 1062

FT MISC-difference 1118

FT MISC-difference 1151

FT MISC-difference 1163

```

ET      /note= "Wild-type Phe substituted by Ser"
FT      Misc-difference 1277
FT      /note= "Wild-type Val substituted by Ile"
FT      Misc-difference 1500
FT      /note= "Wild-type His substituted by Tyr"
FT      Misc-difference 1805
FT      /note= "Wild-type Asp substituted by Asn"
FT      Misc-difference 1930
FT      /note= "Wild-type Ser substituted by Thr"
XX
XX      US6423318-B1.
XX
XX      23-JUL-2002.
XX
XX      31-AUG-2000; 2000US-0653499.
XX
XX      07-JUN-1995; 95US-0475886.
XX      17-SEP-1993; 93US-0397232.
XX      17-SEP-1993; 93WO-US08610.
PR
PR      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PR      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX      WPI: 2002-680946/73.
XX      N-PSDB; ABS52788.
DR
DR      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT      in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX      Example 3: Column 67-78; 71pp: English.
PS
XX      The invention relates to a polynucleotide which encodes a hepatitis A
CC      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC      line). The polynucleotide is useful for preparing a vaccine against
CC      hepatitis A virus infection. This sequence represents a hepatitis A virus
CC      mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
XX      Sequence 2227 AA;
SQ
Query Match 100.0%; Score 107; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 SHIECRKPYKELRLEVGR 20
       |||
       792 SHIECRKPYKELRLEVGR 811

```

Search completed: October 1, 2003, 09:56:40
 Job time : 45.5294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 seconds
(without alignments)
149,988 Million cell updates/sec

Title: US-09-171-432a-39
Perfect score: 107
Sequence: 1 SHIECRKPKYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
al number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pIR1:*
2: pIR2:*
3: pIR3:*
4: pIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	56	2 P00431	genome polyprotein
2	107	100.0	341	2 S04137	genome polyprotein
3	107	100.0	852	1 GNNYHA	genome polyprotein
4	107	100.0	1358	2 A03905	genome polyprotein
5	107	100.0	2227	1 GNNYHM	genome polyprotein
6	107	100.0	2227	1 GNNYHR	genome polyprotein
7	107	100.0	2227	1 GNNYMK	genome polyprotein
8	107	100.0	2227	1 GNNYHB	genome polyprotein
9	105	98.1	56	2 P00427	genome polyprotein
10	104	97.2	56	2 P00429	genome polyprotein
11	102	95.3	319	2 JH0135	genome polyprotein
12	100	93.5	56	2 P00430	genome polyprotein
13	97	90.7	56	2 P00432	genome polyprotein
14	96	89.7	56	2 P00428	genome polyprotein
15	95	88.8	56	2 P00434	genome polyprotein
16	87	81.3	55	2 P00433	genome polyprotein
17	86	80.4	56	2 P00436	genome polyprotein
18	86	80.4	2230	1 GNNYSA	genome polyprotein
19	84.5	79.0	55	2 P00435	genome polyprotein
20	84.5	79.0	839	1 GNNYS2	genome polyprotein
21	81	75.7	56	2 P00437	genome polyprotein
22	46	43.0	420	2 S62541	genome polyprotein
23	44	41.1	374	2 C84040	hypothetical prote
24	43.5	40.7	1663	2 C3KRT	hypothetical prote
25	43	40.2	332	2 P96568	complement C3 prec
26	43	40.2	339	2 T18926	probable lipase, 2
27	43	40.2	343	2 A11823	hypothetical prote
28	43	40.2	346	2 T27896	30S ribosomal prot
29	43	40.2	391	1 VHNZ3	hypothetical prote
					nucleocapsid prote

30	43	40.2	554	2 T49917	hypothetical prote
31	42.5	39.7	140	2 E64364	hypothetical prote
32	42	39.3	121	2 T17708	hypothetical prote
33	42	39.3	214	2 E82679	chaperone XP1452 l
34	42	39.3	286	2 S07532	pufl II/9-1 protei
35	42	39.3	329	2 T25067	hypothetical prote
36	42	39.3	494	2 S62902	legumin 2 precursor
37	42	39.3	723	2 D71091	methionine-tRNA 11
38	42	39.3	861	2 B84963	DNA topoisomerase
39	42	39.3	1087	2 T16876	hypothetical prote
40	42	39.3	1157	2 T43259	pyruvate (flavodox
41	42	39.3	1270	2 T22615	hypothetical prote
42	41.5	38.8	286	2 H87506	hypothetical prote
43	41.5	38.8	606	2 A12649	phosphogluconate d
44	41.5	38.8	606	2 H97431	phosphogluconate d
45	41	38.3	144	2 H96511	AP2 domain contain

ALIGNMENTS

RESULT 1
P00431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totstuka, A.; Naiman, O.V.; Siegl, G.; Wl
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00431
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 107; DB 2; Length 56;
Best local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
Db 29 SHIECRKPKYKELRLEVGKQR 48
|||||

RESULT 2

S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir
A:Reference number: S04137; MUID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein ID (VP1) #status predicted <ANT>

Query Match 100.0%; Score 107; DB 2; Length 341;
Best local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECRKPKYKELRLEVGKQR 20
|||||

Db 296 SHIECRKPYKELRLEVGKOR 315

RESULT 3

GNNYHM

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03904

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.

J. Virol. 54, 247-255, 1985

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MUID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593

Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 107; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 6.3e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 SHIECRKPYKELRLEVGKOR 20

792 SHIECRKPYKELRLEVGKOR 811

RESULT 4

GNNYHM

genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Titchhurst, J.R.; Mele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstone

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA

A:Reference number: A03905; MUID:85186289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 107; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 1e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 SHIECRKPYKELRLEVGKOR 20

792 SHIECRKPYKELRLEVGKOR 811

RESULT 5

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Titchhurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M4707; NID:g329582; PIDN:AAA45465.1; PID:g329583

Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP1>

F:24-243/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP4>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 SHIECRKPYKELRLEVGKOR 20

792 SHIECRKPYKELRLEVGKOR 811

RESULT 6

GNNYHM

genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core

NA polymerase (EC 2.7.7.48); protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03903

R:Naftarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127

A:Accession: A03903

A:Molecule type: genomic RNA

A:Residues: 1-2227 <NAJ>

A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597

Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-980/Product: coat protein 2A #status predicted <C2A>

F:981-1076/Product: core protein 2B #status predicted <C2B>

F:1077-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1507/Product: protein 3B #status predicted <C3B>

F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 SHIECRKPYKELRLEVGKOR 20

792 SHIECRKPYKELRLEVGKOR 811

RESULT 7

GNNYMK
 genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C
 NA polymerase (EC 2.7.7.48), protein 3D
 C:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1988 #sequence, revision 30-Jun-1988 #text-change 16-Jul-1999
 C:Accession: A94149; A23514; A94508
 R:Coen, J. I.; Rosenblum, B.; Ticehurst, J. R.; Daemer, R. J.; Feinstein, S. M.; Purcell, R. H.; Proctor, Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
 A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
 A:Reference number: A94149; M01D:81175701; PMID:3031686
 A:Accession: A94149
 A:Status: nucleic acid sequence not shown
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <COH>
 A:Cross-references: EMBL:M16632; M1D:g329594; P1D:NAAA45471.1; P1D:g329595
 A:Note: Submitted to Genbank, August 1987
 C:Superfamily: hepatitis A virus genome polypeptide
 C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
 F:1-245/Product: coat protein 1A #status predicted <P1A>
 F:46-491/Product: coat protein 1B #status predicted <P1B>
 F:492-836/Product: coat protein 1C #status predicted <P1C>
 F:837-980/Product: core protein 2A #status predicted <P2A>
 F:981-1076/Product: core protein 2B #status predicted <P2B>
 F:1077-1423/Product: core protein 2C #status predicted <P2C>
 F:1423-1484/Product: protein 3A #status predicted <P3A>
 F:1485-1507/Product: protein 3B #status predicted <P3B>
 F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
 F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

```

Query Match          100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;
Matches    20; Conservative   0; Mismatches     0; Indels      0; Gaps      0;

Oy 1 SHIECRPKYELRLEVGKOR 20
|||||
Db 792 SHIECRPKYELRLEVGKOR 811

RESULT 8
GNMYHB
genome polyprotein - human hepatitis A virus (strain MBF)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
Accession: J50303
Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhardt,
virus Res. 8, 153-171, 1987
A>Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <P&U>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydro
F:1-23/Product: coat protein 1A #status predicted <VP>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1518/Product: genome-linked protein VPg #status predicted <VPG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match          100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;

```

	Matches	20; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1	SHIECRPYKELRELVCKQR	20		
Db	792	SHIECRPYKELRELVCKQR	811		

```

RESULT 9
P00427
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Slegl, G.; W
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00427
A:Molecule type: RNA
A:Residues: 1-56 <R0B>
C:Comment: This protein is from the VP1/2A Junction region.
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match          98.18; Score 105; DB 2; Length 56;
Best Local Similarity 95.08; Pred. No. 8.2e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

0y      1 SHIECRKPYKELRLEVQKOR 20
Db      | | | | | | | | | | | | | |
        29 SHIECRKPYKELRLEVQKOR 48

```

RESULT 10
PQ0429
genome polyprotein - human hepatitis A virus (strain PRCl6) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Slegl, G.; W
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match	97.28;	Score 104;	DB 2;	Length 56;
Best Local Similarity	95.08;	Pred. NO. 1.2e-09;		
Matches 19; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      1 SHIECRKPYKELRLEVGKQR 20
          |||||:|||||
Db      29 SHIECRKPYKDLRLEVGKQR 48

```

RESULT 11
JH0135
genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence-revision 16-Sep-1992 #text-change 08-Apr-1994
C:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
virus Res. 6, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus
A:Reference number: JH0135; MUID:88129044; PMID:2829458
A:Accession: JH0135
A:Molecule type: genomic RNA

A;Residues: 1-319 <ROB>
A;Cross-references: GB:M22821
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-300/Product: coat protein ID #status predicted <CPD>
F:301-319/Product: core protein 2A (fragment) #status predicted <C2A>
F:237/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 95.3%; Score 102; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 19
|||||
Db 301 SHIECRKPYKELRLVGVKOR 319

RESULT 12

PQ0430 genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)

C;Accession: PQ0430
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0430
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 93.5%; Score 100; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 5.2e-09;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20
|||||
Db 29 SHIECRKPYKELRLVGVKOR 48

RESULT 13

PQ0432 genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)

C;Accession: PQ0432
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0432
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 90.7%; Score 97; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.6e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20
|||||
Db 29 SHIECRKPYKELRLVGVKOR 48

RESULT 14

PQ0428 genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C;Species: human hepatitis A virus

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0428
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0428
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 89.7%; Score 96; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 2.3e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20
|||||
Db 29 SHIECRKPYKELRLVGVKOR 48

RESULT 15

PQ0434 genome polyprotein - human hepatitis A virus (strain KPH) (fragment)

C;Accession: PQ0434
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0434
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 88.8%; Score 95; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 3.3e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20
|||||
Db 29 SHIECRKPYKELRLVGVKOR 48

Search completed: October 1, 2003, 10:04:33
Job time : 15.8235 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds
(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-39

Perfect score: 107
Sequence: 1 SHIECRKPYKELRLEVGKOR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_41:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	341	1	POLG_HPAV1
2	107	100.0	852	1	POLG_HPAVC
3	107	100.0	2226	1	POLG_HPAV2
4	107	100.0	2226	1	POLG_HPAV4
5	107	100.0	2226	1	POLG_HPAV8
6	107	100.0	2227	1	POLG_HPAV8
7	107	100.0	2227	1	POLG_HPAV1
8	107	100.0	2227	1	POLG_HPAV1
9	95	88.8	808	1	POLG_HPAV1
10	86	80.4	2230	1	POLG_HPAV1
11	84.5	79.0	839	1	POLG_HPAV1
12	46	42.0	420	1	YAGA_SCHPO
13	45	42.1	1819	1	CGP6_HUMAN
14	43.5	40.7	1663	1	CO3_RAT
15	43	40.2	391	1	NCAP_HRSV1
16	42	39.3	210	1	LOLA_XYLEA
17	42	39.3	286	1	PU91_SCICO
18	42	39.3	723	1	SYM_PYRHO
19	42	39.3	861	1	TOP1_BUCAI
20	41	38.3	273	1	R12_BUCAP
21	41	38.3	298	1	EMD4_BACHD
22	41	38.3	319	1	URED_SYNPV
23	41	38.3	364	1	SUCG_METVA
24	41	38.3	370	1	T2M2_METVA
25	41	38.3	391	1	NCAP_BRSV3
26	41	38.3	391	1	NCAP_BRSV3
27	41	38.3	391	1	NCAP_BRSV3
28	41	38.3	391	1	NCAP_BRSV3
29	41	38.3	492	1	ANKH_HUMAN
30	41	38.3	492	1	ANKH_MOUSE
31	41	38.3	492	1	ANKH_MOUSE
32	41	38.3	492	1	ANKH_MOUSE
33	41	38.3	840	1	NI61_HUMAN

34	41	38.3	843	1	NI61_MOUSE	Q99K10 mus musculus
35	41	38.3	843	1	NI61_RAT	Q62765 rattus norv
36	41	38.3	955	1	T150_HUMAN	O9Y2W1 homo sapien
37	41	38.3	4543	1	LRP1_CHICK	P98157 gallus galli
38	41	38.3	4544	1	LRP1_HUMAN	O07954 homo sapien
39	40.5	37.9	141	1	V16K_TRTVC	P05076 tobacco rat
40	40.5	37.9	274	1	NAGB_FUSNN	O8REG1 fusobacteri
41	40.5	37.9	1120	1	RPOW_SCHPO	O13993 schizosacch
42	40.5	37.9	1663	1	CO3_MOUSE	P01027 mus musculus
43	40.5	37.9	2198	1	YLJ2_CAEEL	P34367 caenorhabdi
44	40	37.4	69	1	MOP_HAEIN	P45183 haemophilus
45	40	37.4	238	1	PTH_AGRTS	Q8UD97 agrobacteri

ALIGNMENTS

RESULT 1
POLG_HPAV1 STANDARD: PRT; 341 AA.
AC P13672:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP3; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain LCDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV)."
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -I- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14666; GAA32794.1; .
DR PIR: S04137; S04137.
KW Polypeptide; Coat protein; Core protein.
FT NON_TER 1
FT CHAIN 1
FT CHAIN 2 340 COAT PROTEIN VP3 (1C).
FT CHAIN 341 >341 COAT PROTEIN VP1 (1D).
FT NON_TER 341 341 CORE PROTEIN P2A.
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
Query Match 100.0%; Score 107; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELRLEVGKOR 20
DB 296 SHIECRKPYKELRLEVGKOR 315
RESULT 2
POLG_HPAVC STANDARD: PRT; 852 AA.
AC P06442; Q83741; Q83742;

DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12097;
RX MEDLINE=915648; PubMed=2985793;
RA Linemeyer D.L., Wenke J.G., Martin-Gallardo A., Hughes J.V., Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA"; J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M10033; AAA45470.1; -.
DR PIR: A03904; GNNYHA.
KW Polypeptide; Coat protein; Core protein.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 >852
FT NON_TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3BD0AD532820E CRC64;
Query Match 100.0%; Score 107; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKOR 20
Db 792 SHIECRKPKYKELRLVGVKOR 811
RESULT 3
ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48).
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12094;
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Crommons T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination"; J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + RNA(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59810; AAA45468.1; -.
DR MEROPS; C03.005; -.
DR InterPro: IPR004004; Calic_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSV1r.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
Query Match 100.0%; Score 107; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKOR 20
Db 792 SHIECRKPKYKELRLVGVKOR 811
RESULT 4
ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48).
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12095;
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Crommons T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination"; J. Virol. 65:2056-2065(1991).

```
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59809; AAA45469.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol: 1.
DR Pfam: PF00910; RNA_helicase: 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGRKOR 20
DB 792 SHIECRKPKYKELRLEVGRKOR 811

RESULT 5
POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID POLG_HPAV8
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cromeans T., Jansen R.W.;
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RA variants arising during persistent infection: evidence for genetic
RA recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
```

```
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59808; AAA45467.1; -.
DR PDB: 1OAT; 15-MAY-00.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol: 1.
DR Pfam: PF00910; RNA_helicase: 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGRKOR 20
DB 792 SHIECRKPKYKELRLEVGRKOR 811

RESULT 6
POLG_HPAV8 STANDARD; PRT; 2227 AA.
ID POLG_HPAV8
AC P08617; P06443; O81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RA "Complete nucleotide sequence of wild-type hepatitis A virus:
RA comparison with different strains of hepatitis A virus and other
RA picornaviruses.";
RL J. Virol. 61:50-59(1987).
CC [2]
CC SEQUENCE FROM N.A.
```

RC STRAIN-Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.T., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RT Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RL comparison with wild-type virus.";
RN Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RP [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maize J.V. Jr.,
RT Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RL proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate. +
CC (NNA)(N).
CC
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC
CC -I- PRIM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -I- MISCELLANEOUS: STRAIN HM-175/77 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -I- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between The Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isdb.ch/announce/>
CC or send an email to license@isdb.slb.ch).
CC -----

DR EMBL; M14114; AAA45475.1; -;
DR EMBL; M14707; AAA45465.1; -;
DR EMBL; M14707; AAA45466.1; ALT_INTT.
DR EMBL; M16632; AAA45471.1; -;
DR PIR; A25981; GNNYTM.
DR PIR; A94149; GNNYTM.
DR PDB; 1HAV; 23-DEC-96.
DR MEROPS; C03.005; -;
DR InterPro; IPR004004; Calic.pol.bel.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_PSD.
DR InterPro; IPR007094; RNA_pol_PSVtr.
Pfam; PF00680; RNA_dep_RNA_pol_1.
PFam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICYRUSNS.

KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT CHAIN 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277

COAT PROTEIN VP4 (PIA).
COAT PROTEIN VP2 (PIB).
COAT PROTEIN VP3 (PIC).
COAT PROTEIN VP1 (PID).
CORE PROTEIN P2A.
CORE PROTEIN P2B.
CORE PROTEIN P2C.
PROBABLE PROTEIN P3A.
PROBABLE PROTEIN P3B.
PROBABLE PROTEIN P3C.
RNA-DIRECTED POLYMERASE 3D.
K->R (IN ATTENUATED STRAIN).
E->V (IN ATTENUATED STRAIN).
N->S (IN ATTENUATED STRAIN).
A->V (IN ATTENUATED STRAIN).
G->A (IN ATTENUATED STRAIN).
K->M (IN ATTENUATED STRAIN).
E->K (IN ATTENUATED STRAIN).
F->S (IN ATTENUATED STRAIN).
V->I (IN ATTENUATED STRAIN).

FT	VARIANT	1500	1500	H -> Y (IN ATTENUATED STRAIN).
FT	VARIANT	1805	1805	D -> N (IN ATTENUATED STRAIN).
FT	VARIANT	1930	1930	S -> T (IN ATTENUATED STRAIN).
SO	SEQUENCE	2227 AA.	251506 MW;	016225E7AEB740A6 CRC64;
Query Match				
Best Local Similarity		100.0%;	Score 107;	DB 1; Length 2227;
Matches	20;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 SHIECRKPKYLEVLEVGKOR 20 			
Db	792 SHIECRKPKYLEVLEVGKOR 811			
RESULT 7				
POLG_HPAVL	STANDARD;	PRT;	2227 AA.	
ID				
AC	P06441.			
DT	01-JAN-1988 (Rel. 06, Created)			
DT	01-JAN-1988 (Rel. 06, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polypeptin [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (PC 2.7.7.48)].			
DE	P3D (PC 2.7.7.48)].			
OS	Hepatitis A virus (strain LA),			
OC	Virusess; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
CC	Hepatovirus.			
OX	NCBI_Taxid=12099;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=65190549; PubMed=2986127;			
RA	Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,			
RA	Merrweather J., van Nest G., Dina D.;			
RT	Primary structure and gene organization of human hepatitis A virus."/;			
RL	Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985) .			
CC	-1- CAPSITIC ACTIVITY: N nucleoside triphosphate +			
CC	(RNA)(N).			
CC	-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,			
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,			
CC	VP3, AND VP4.			
CC	-1- PRIM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between The Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isdb-sib.ch/announce/			
CC	or send an email to license@isdb-sib.ch).			
CC	-----			
DR	EMBL; K02990; AAA45472.1; -.			
DR	PIR; A03903; GNMYHR.			
DR	MEROPS; CO3_005; -.			
DR	InterPro; IPR004004; Calic1.pol.hel.			
DR	InterPro; IPR000605; RNA_helicase.			
DR	InterPro; IPR007095; RNA_pol_DS_PS.			
DR	InterPro; IPR001205; RNA_pol_P3D.			
DR	InterPro; IPR007094; RNA_pol_PSVr.			
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.			
DR	Pfam; PF00910; RNA_helicase; 1.			
DR	PRINTS; PR00918; CALICIVIRUSN.			
KW	Polyprotein; Coat protein; Core protein; Transferase;			
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.			
FT	CHAIN	1	23	COAT PROTEIN VP4 (P1A).
FT	CHAIN	24	245	COAT PROTEIN VP2 (P1B).
FT	CHAIN	246	491	COAT PROTEIN VP3 (P1C).
FT	CHAIN	492	836	COAT PROTEIN VP1 (P1D).
FT	CHAIN	837	980	CORE PROTEIN P2A.
FT	CHAIN	981	1076	CORE PROTEIN P2B.
FT	CHAIN	1077	1422	CORE PROTEIN P2C.
FT	CHAIN	1423	1484	PROBABLE PROTEIN P3A.
FT	CHAIN	1485	1507	PROBABLE PROTEIN P3B.

FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
 FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2227 AA: 251898 MW: 99A7354B4CD2799C CRC64:
 Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGNOR 20
 |||||
 DB 792 SHIECRKPKYKELRLEVGNOR 811

RESULT 8
 POLG_HPAVM STANDARD; PRT: 2227 AA.
 AC P13901: Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (rel. 13, Created)
 DT 01-JAN-1990 (rel. 13, Last sequence update)
 DE 28-FEB-2003 (rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MBB).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; Pubmed-2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Klein R., Wimmer E.,
 RA Delnhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB)."
 RT Virus Res 8:153-171(1987).
 RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@sdb-sib.ch).
 CC -----
 CC EMBL: M20273; AAA45474.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P1D.
 DR InterPro: IPR007094; RNA_pol_PSVT.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KM Polyprotein: Coat protein; Core protein; Transferase;
 KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 980 CORE PROTEIN P2A.
 FT CHAIN 981 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.

FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2227 AA: 251425 MW: EC983ED2A7C86349 CRC64:
 Query Match 100.0%; Score 107; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGNOR 20
 |||||
 DB 792 SHIECRKPKYKELRLEVGNOR 811

RESULT 9
 POLG_HPAVM STANDARD; PRT: 808 AA.
 AC Q02381;
 DT 01-JUL-1993 (rel. 26, Created)
 DT 01-JUL-1993 (rel. 26, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 OS Hepatitis A virus (strain GA76).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=31706;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92260183; Pubmed-1316423;
 RA Khanna B., Speilbring J.E., Innis B.L., Robertson B.H.;
 RT "Characterization of a genetic variant of human hepatitis A virus."
 RT J. Med. Virol. 36:118-124(1992).
 RL -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to license@sdb-sib.ch).
 CC -----
 CC EMBL: M6695; AAA45477.1; -
 DR Polyprotein: coat protein; Core protein.
 KM NON_TER 1 1
 FT CHAIN 1 2 COAT PROTEIN VP4 (P1A).
 FT CHAIN 3 223 COAT PROTEIN VP2 (P1B).
 FT CHAIN 224 470 COAT PROTEIN VP3 (P1C).
 FT CHAIN 471 770 COAT PROTEIN VP1 (P1D).
 FT CHAIN 771 >808 CORE PROTEIN P2A.
 FT NON_TER 808 808
 SQ SEQUENCE 808 AA: 90632 MW: D80CE7E57A479C12 CRC64:
 Query Match 88.8%; Score 95; DB 1; Length 808;
 Best Local Similarity 95.0%; Pred. No. 1.5e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGNOR 20
 |||||
 DB 771 SHIECRKPKYKELRLEVGNOR 790

RESULT 10
 POLG_HPAVM STANDARD; PRT: 2230 AA.
 AC P14553;
 DT 01-JAN-1990 (rel. 13, Created)
 DT 01-AUG-1992 (rel. 23, Last sequence update)
 DT 28-FEB-2003 (rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins


```

DE P2A TO P2C; probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27) .
OC viruses; ssRNA positive strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OK NCBI_TaxID=12102;
RX SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Tichehurst J.R.,
RA Purrell R.H.;
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RA structure and growth in cell culture with other HAV strains.";
RA J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapaidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RA "Variations in genome fragments coding for RNA polymerase in human
RA and simian hepatitis A viruses.";
RA PEBs Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PFM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/issb-sdb.ch).
CC or send an email to license@issb-sdb.ch.
CC -----
CC EMBL, D00924; BAA00766.1; -
CC EMBL, X15461; CA33490.1; -
CC PIR, A30470; GNNYSA.
DR MEROPS: C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_Pf.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSV1r.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR Polyprotein; Coat protein; Core protein; Transferrase;
DR RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
KM CHAIN 1 27
FT CHAIN 28 249 COAT PROTEIN VP4 (PIA).
FT CHAIN 250 495 COAT PROTEIN VP2 (PIB).
FT CHAIN 496 795 COAT PROTEIN VP3 (PIC).
FT CHAIN 796 984 COAT PROTEIN VP1 (PID).
FT CHAIN 985 1091 CORE PROTEIN P2A.
FT CHAIN 1092 1426 CORE PROTEIN P2B.
FT CHAIN 1427 1498 CORE PROTEIN P2C.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3A.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3B.
FT CHAIN 1742 2230 PROBABLE PROTEIN 3C.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3320E324E1F19 CRC64;

Query Match 80.4%; Score 86; DB 1; Length 2230;
Best local Similarity 85.0%; Pred. No. 1.2e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
1 SHIECRKPKYKELRLEVGRKOR 20
||||| |||||||:|||||
796 SHIEGRKPKYKELRLEVGRKOR 815

```

```

RESULT 11
POLG_HPAVT STANDARD: PRT: 839 AA.
ID AC P31788;
DC 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein
P2a] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91311421; PubMed=1649902;
RA Nathan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1 SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1 PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M59286; AAA5473.1; -.
DR PIR: J01180; GNNS2.
KW Polypeptide; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2ACCA4BD1E192DBC CRC64;
Query Match 79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 7,7e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 SHIECRPKYKELREVGKOR 20
Db 792 SHIECRPKYKELREVGKOR 810
IIII:IIIIIIIIIIIIII
RESULT 12
YAGA_SCHPO STANDARD: PRT: 420 AA.
ID AC Q09873;
DC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C12G12.10 in chromosome 1.
GN SPAC12G12.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

```


RA Squire J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gough S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Moore P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoprie B.,
 RA Wajsbom I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Filtz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Mambutti R., Punelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaive V., Motter S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerruti L., Love T., McCombie M.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe".
 RL Nature 415:871-880(2002).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 CC EMBL: 265568; CAA91505.1; -
 DR PIR: S62541; S62541.
 DR GeneDB:SPombe:SPAC12G12.10; -
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 1.
 DR SMART: SM00320; WD40; 3.
 KW Hypothetical protein.
 SQ SEQUENCE 420 AA; 47525 MW; 9DC1F710FBEB9761 CRC64;
 Query Match 43.0%; Score 46; DB 1; Length 420;
 Best Local Similarity 46.7%; Pred. No. 6.2;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 4 ECRPKYKELRELVGK 18
 376 DCSLPFKEMRYDDGK 390
 RESULT 13
 ID GCP6_HUMAN STANDARD; PRT; 1819 AA.
 AC Q96RT7; Q9BY91; Q9UGX3; Q9UGX4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gamma-tubulin complex component 6 (GCP-6).
 GN TUBGCP6 OR GCP6 OR KIAA1659.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1) AND CHARACTERIZATION.
 RX MEDLINE=21551508; PubMed=11694571.
 RA Murphy S.M., Preble A.M., Patel U.K., O'Connell K.L., Dias D.P.,
 RA Murphy M., Agard D., Stuifjes J.T., Stearns T.,
 RT "GCP5 and GCP6: two new members of the human gamma-tubulin complex".
 RL Mol. Biol. Cell 12:3340-3352(2001).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasly O.P.,
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fay J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafton D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Famlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leverisha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
 RA Mcclay J., McLaren S., McMurtry A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Steward C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Corcos M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Peplin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
 RA Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
 RA Kim U.J., Shizuya H., Simon M.I., Dumaniski J.P., Peyraud M., Kedra D.,
 RA Seroussi E., Fransson I., Tapia I., Brudner C.F., O'Brien K.P.,
 RA Wilkinson P., Bodencelch A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tiliun Y., Wright H.,
 RT "The DNA sequence of human chromosome 22".
 RL Nature 402:489-495(1999).
 [3]
 RP SEQUENCE OF 33-1819 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=21156230; PubMed=11258795;
 RA Hirotsawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.,
 RT "Identification of novel transcribed sequences on human chromosome 22
 RT by expressed sequence tag mapping.";
 RL DNA Res. 8:1-9(2001).
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule
 CC nucleation at the centrosome.
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
 CC GCP3, GCP4, GCP5 and GCP6.
 CC -1- SUBCELLULAR LOCATION: Centrosome.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q96RT7-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q96RT7-2; Sequence=VSP_001624;
 CC Note=No experimental confirmation available;
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
 CC in positions 1371 and 1758.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions in no way
 CC use by non-profit institutions as long as its content is in no way

FT CONFLICT 721 722 LK -> KL (IN REF. 2).
 SQ SEQUENCE 1663 AA; 186460 MW; 2F87C8143CDD4BC CRC64;

Query Match 40.7%; Score 43.5; DB 1; Length 1663;
 Best Local Similarity 47.4%; Pred. NO. 63;

Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

OY 1 SHIECRKPYKELRLVGRQ 19
 ||::|| |::|||
 Db 1586 SHVACR---NALKLOKCKQ 1601

RESULT 15

NCAP_HRSV1 STANDARD; PRT: 391 AA.

AC P24566;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 Nucleocapsid protein.

N.

Human respiratory syncytial virus (subgroup B / strain 18537).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI_TaxID=11251;

RP [1]

RP MEDLINE=89279331; PubMed=2525176;

RA Johnson P.R., Collins P.L.;

RT "The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial

RT virus (RSV) of antigenic subgroups A and B: sequence conservation and

RT divergence within RSV genomic RNA.";

RL J. Gen. Virol. 70:1539-1547 (1989).

-1- FUNCTION: MOST ABUNDANT PROTEIN IN THE VIRION AND AN IMPORTANT

ELEMENT CONFERRING HELICAL SYMMETRY ON THE NUCLEOPROTEIN CORE AS

WELL AS INTERACTING WITH M PROTEIN DURING VIRION FORMATION.

-1- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES NUCLEOCAPSID FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL; D00736; BAA00637.1; -

DR PIR; G32063; VHN23.

DR InterPro; IPR004930; Pneumo_ncap.

DR Pfam; PF03246; Pneumo_ncap.1.

DR Prodom; PD006438; Pneumo_ncap.1.

DR Nucleocapsid
 SO SEQUENCE 391 AA; 43416 MW; 9A2BBB50103B2835 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 391;

Best Local Similarity 56.2%; Pred. NO. 17;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 IECKRPYKELRLVGR 18
 |||||::|
 Db 129 IESRSYKRLKEMGE 144

Search completed: October 1, 2003, 09:57:45
 Job time : 9.58824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107
Sequence: 1 SHIECRKPKYKELRLEVGKOR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	53	12	Q98VY2 hepatitis a
2	107	100.0	53	12	Q98VY3 hepatitis a
3	107	100.0	53	12	Q98VY4 hepatitis a
4	107	100.0	53	12	Q98VY0 hepatitis a
5	107	100.0	53	12	Q98VY5 hepatitis a
6	107	100.0	53	12	Q98VY6 hepatitis a
7	107	100.0	53	12	Q98VY1 hepatitis a
8	107	100.0	53	12	Q98VY7 hepatitis a
9	107	100.0	55	12	Q98VY2 hepatitis a
10	107	100.0	55	12	Q98VY4 hepatitis a
11	107	100.0	55	12	Q98VY3 hepatitis a
12	107	100.0	55	12	Q98VY0 hepatitis a
13	107	100.0	55	12	Q98VY5 hepatitis a
14	107	100.0	56	12	Q98VY6 hepatitis a
15	107	100.0	56	12	Q98VY1 hepatitis a
16	107	100.0	56	12	Q98VY7 hepatitis a

17	107	100.0	56	12	Q98VY8 hepatitis a
18	107	100.0	56	12	Q98VY2 hepatitis a
19	107	100.0	56	12	Q98VY3 hepatitis a
20	107	100.0	56	12	Q98VY1 hepatitis a
21	107	100.0	56	12	Q98VY0 hepatitis a
22	107	100.0	56	12	Q98VY5 hepatitis a
23	107	100.0	56	12	Q98VY6 hepatitis a
24	107	100.0	56	12	Q98VY1 hepatitis a
25	107	100.0	56	12	Q98VY7 hepatitis a
26	107	100.0	56	12	Q98VY2 hepatitis a
27	107	100.0	56	12	Q98VY4 hepatitis a
28	107	100.0	56	12	Q98VY3 hepatitis a
29	107	100.0	56	12	Q98VY0 hepatitis a
30	107	100.0	56	12	Q98VY5 hepatitis a
31	107	100.0	56	12	Q98VY6 hepatitis a
32	107	100.0	56	12	Q98VY1 hepatitis a
33	107	100.0	56	12	Q98VY7 hepatitis a
34	107	100.0	56	12	Q98VY2 hepatitis a
35	107	100.0	56	12	Q98VY4 hepatitis a
36	107	100.0	56	12	Q98VY3 hepatitis a
37	107	100.0	56	12	Q98VY0 hepatitis a
38	107	100.0	56	12	Q98VY5 hepatitis a
39	107	100.0	56	12	Q98VY6 hepatitis a
40	107	100.0	56	12	Q98VY1 hepatitis a
41	107	100.0	56	12	Q98VY7 hepatitis a
42	107	100.0	56	12	Q98VY2 hepatitis a
43	107	100.0	56	12	Q98VY4 hepatitis a
44	107	100.0	56	12	Q98VY3 hepatitis a
45	107	100.0	56	12	Q98VY0 hepatitis a

ALIGNMENTS

RESULT 1

Q98VY2	PRELIMINARY;	PRT;	53 AA.
AC Q98VY2;			
DT 01-JUN-2001 (TREMBLrel. 17, Created)			
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE Polypeptide (Fragment).			
OS Hepatitis A virus.			
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC Hepatovirus.			
OX NCBI_TaxID=12092;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN-Uruguay6;			
RA Costa-Matloff M., Ferre V., Monpocho S., Garcia L., Collina R.,			
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;			
RT "Genetic variability of Hepatitis A virus in South America reveals			
RT heterogeneity and co-circulation during epidemic outbreaks."			
RL Submitted (Apr-2001) to the EMBL/Genbank/DBD databases.			
DR EMBL; AJ309232; CAC37078.1; -			
FT NON_TER	1	1	
FT NON_TER	53	53	
SO SEQUENCE	53 AA;	6271 MW;	0846AEFABC397432 CRC64;
Query Match	100.0%;	Score 107;	DB 12;
Best Local Similarity	100.0%;	Pred. No. 4.2e-10;	Length 53;
Matches	20;	Conservative	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;
QY 1 SHIECRKPKYKELRLEVGKOR 20			
DB 29 SHIECRKPKYKELRLEVGKOR 48			
RESULT 2			
Q98VY3	PRELIMINARY;	PRT;	53 AA.
AC Q98VY3;			
DT 01-JUN-2001 (TREMBLrel. 17, Created)			

DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RN NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Drul6;
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,
RT "Genetic variability of Hepatitis A virus in South America reveals
heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309231; CAC37077.1; -.
FT NON_TER 1 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;
Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20
DB 29 SHIECRKPYKELRLEVGKOR 48
|||||

RESULT 3
Q98YV4 PRELIMINARY; PRT; 53 AA.
AC Q98YV4;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RN NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-J;
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,
RT "Genetic variability of Hepatitis A virus in South America reveals
heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309230; CAC37076.1; -.
FT NON_TER 1 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20
DB 29 SHIECRKPYKELRLEVGKOR 48
|||||

RESULT 4
Q98YV0 PRELIMINARY; PRT; 53 AA.
AC Q98YV0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Uruguay4;
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,
RT "Genetic variability of Hepatitis A virus in South America reveals
heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309234; CAC37080.1; -.
FT NON_TER 1 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;
Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20
DB 29 SHIECRKPYKELRLEVGKOR 48
|||||

RESULT 5
Q98YV5 PRELIMINARY; PRT; 53 AA.
AC Q98YV5;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RN NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Drul3;
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,
RT "Genetic variability of hepatitis A virus in South America reveals
heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309229; CAC37075.1; -.
FT NON_TER 1 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 4.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20
DB 29 SHIECRKPYKELRLEVGKOR 48
|||||

RESULT 6
Q98YV6 PRELIMINARY; PRT; 53 AA.
AC Q98YV6;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DE 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RN NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Drul7;

RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
 RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
 RT "Genetic variability of hepatitis A virus in South America reveals
 RT heterogeneity and co-circulation during epidemic outbreaks.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ309228; CAC37074.1; -.
 FT NON_TER 1
 FT NON_TER 53
 SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
 Best Local Similarity 100.0%; Pred. No. 4.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20
 DB 29 SHIECRKPKYKELREVGKOR 48

RESULT 7

O98VY1 PRELIMINARY; PRT; 53 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.

NCBI_TaxID=12092;

RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;

RT "Genetic variability of Hepatitis A virus in South America reveals
 RT heterogeneity and co-circulation during epidemic outbreaks.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ309233; CAC37079.1; -.

FT NON_TER 1
 FT NON_TER 53
 SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
 Best Local Similarity 100.0%; Pred. No. 4.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SHIECRKPKYKELREVGKOR 20
 29 SHIECRKPKYKELREVGKOR 48

RESULT 8

O98VY7 PRELIMINARY; PRT; 53 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.

NCBI_TaxID=12092;

RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;

RT "Genetic variability of hepatitis A virus in South America reveals
 RT heterogeneity and co-circulation during epidemic outbreaks.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ309227; CAC37073.1; -.
 FT NON_TER 1
 FT NON_TER 53
 SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
 Best Local Similarity 100.0%; Pred. No. 4.2e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20
 DB 29 SHIECRKPKYKELREVGKOR 48

RESULT 9

O999T2 PRELIMINARY; PRT; 55 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.

NCBI_TaxID=12092;

RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;

RT "Genetic variability of Hepatitis C virus reveals heterogeneity and
 RT co-circulation during epidemic outbreaks.";
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ306386; CAC29235.1; -.

FT NON_TER 1
 FT NON_TER 55
 SQ SEQUENCE 55 AA: 6513 MW: 135CH05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;
 Best Local Similarity 100.0%; Pred. No. 4.3e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20
 DB 28 SHIECRKPKYKELREVGKOR 47

RESULT 10

O999T4 PRELIMINARY; PRT; 55 AA.

DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.

NCBI_TaxID=12092;

RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;

RT "Genetic variability of Hepatitis C virus reveals heterogeneity and
 RT co-circulation during epidemic outbreaks.";
 RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ306384; CAC29233.1; -.
 DR InterPro: IPR000886; ER_TARGET.
 DR PROSITE: PS00014; ER_TARGET; 1.

FT NON_TER 1 1
SQ SEQUENCE 55 AA: 6513 MW: 135CB05D46AEF4CB CRC64;
Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLEVGKOR 20
DB 28 SHIECRKPKYKELRLEVGKOR 47
RESULT 11
ID 099973 PRELIMINARY; PRT; 55 AA.
AC 099973;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DE 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
OS Polypeptide (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Chile-9;
RA Balladon S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306385; CAC29234.1; -;
DR InterPro; IPR000886; ER-target.
DR PROSITE; PS00014; ER-TARGET; 1.
FT NON_TER 1 1
SQ SEQUENCE 55 AA: 6513 MW: 135CB05D46AEF4CB CRC64;
Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLEVGKOR 20
DB 28 SHIECRKPKYKELRLEVGKOR 47
RESULT 12
ID 099908 PRELIMINARY; PRT; 55 AA.
AC 099908;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DE 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
OS Polypeptide (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Arg-6;
RA Balladon S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis A virus in South America reveals heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306370; CAC29219.1; -;
DR InterPro; IPR000886; ER-target.
DR PROSITE; PS00014; ER-TARGET; 1.
FT NON_TER 1 1

FT NON_TER 55 55
SQ SEQUENCE 55 AA: 6513 MW: 135CB05D46AEF4CB CRC64;
Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 4.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLEVGKOR 20
DB 28 SHIECRKPKYKELRLEVGKOR 47
RESULT 13
ID 08794 PRELIMINARY; PRT; 56 AA.
AC 08794;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DE 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Manuvelo;
RA Theamboonlers A., Jantaradassamee P., Poonvorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an outbreak in the southern part of Thailand.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507065; AAM3425.1; -;
DR InterPro; IPR000886; ER-target.
DR PROSITE; PS00014; ER-TARGET; 1.
FT NON_TER 1 1
SQ SEQUENCE 56 AA: 6642 MW: 465CF51846AEF4BC CRC64;
Query Match 100.0%; Score 107; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLEVGKOR 20
DB 29 SHIECRKPKYKELRLEVGKOR 48
RESULT 14
ID 067822 PRELIMINARY; PRT; 56 AA.
AC 067822;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
OS Polypeptide (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ISR-70;
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S., Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77247; CAB01040.1; -;
DR InterPro; IPR000886; ER-target.
DR PROSITE; PS00014; ER-TARGET; 1.
FT NON_TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.

FT NON_TER 56 56
 SO SEQUENCE 56 AA; 6614 MW; 8438C51846A6F4A4 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;
 Best Local Similarity 100.0%; Pred. No. 4,4e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGR 20
 |||||
 Db 29 SHIECRKPYKELRLEVGR 48

RESULT 15

039872 PRELIMINARY; PRT; 56 AA.

ID 039872
 AC 039872;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).

HE Hepatitis A virus.
 Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=VDM;
 RX MEDLINE=97247817; PubMed=9093940;

RA Taylor M.B.;
 RT "Molecular epidemiology of South African strains of hepatitis A virus:
 J. Med. Virol. 51:273-279(1997).

RL EMBL; U68697; AAB53593.1; -
 DR InterPro; IPR000886; ER_target.

DR PROSITE; PS00014; ER_TARGET; 1.

FT NON_TER 1 1

FT NON_TER 56 56

SO SEQUENCE 56 AA; 6642 MW; 465CF51846A6F4BC CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;
 Best Local Similarity 100.0%; Pred. No. 4,4e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGR 20
 |||||
 Db 29 SHIECRKPYKELRLEVGR 48

Search completed: October 1, 2003, 10:02:34
 Time : 35.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds
(without alignments)
66.600 Million cell updates/sec

Title: US-09-171-432a-39

Perfect score: 107
Sequence: 1 SHIECRKPKELRLEVGKOR 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTMUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1091	6	US-08-475-886-2
2	107	100.0	2227	3	US-08-475-886-2
3	107	100.0	2227	3	US-08-475-886-4
4	107	100.0	2227	3	US-08-475-886-6
5	107	100.0	2227	3	US-08-397-232-2
6	107	100.0	2227	3	US-08-397-232-4
7	107	100.0	2227	3	US-09-171-387-2
8	107	100.0	2227	4	US-09-653-499-2
9	107	100.0	2227	4	US-09-653-499-4
10	107	100.0	2227	4	US-09-653-499-6
11	84.5	79.0	839	1	US-08-087-016-2
12	43	40.2	1127	3	US-09-150-460B-11
13	42.5	39.7	5405	3	US-08-718-388-9
14	42	39.3	4444	4	US-09-252-991A-32415
15	42	39.3	472	4	US-09-328-352-6296
16	41	38.3	114	4	US-09-489-847-321
17	41	38.3	391	5	PCR-US91-08177-3
18	41	38.3	534	4	US-09-252-991A-2086
19	41	38.3	843	4	US-09-491-356C-20
20	41	38.3	4544	1	US-08-469-486-52
21	41	38.3	4544	2	US-08-469-486-52
22	40	37.4	110	4	US-09-341-461-21
23	40	37.4	414	4	US-09-252-991A-26659
24	39	36.4	24	1	US-08-406-347A-13
25	39	36.4	154	4	US-09-198-452A-530
26	39	36.4	238	4	US-09-257-179-80
27	39	36.4	305	4	US-09-635-872A-3

ALIGNMENTS

28	39	36.4	305	4	US-09-636-077A-3	Sequence 3, Appl1
29	39	36.4	344	3	US-09-393-554-2	Sequence 2, Appl1
30	39	36.4	495	4	US-09-252-991A-25802	Sequence 25802, A
31	39	36.4	520	6	5223391-5	Patent No. 5223391
32	39	36.4	892	4	US-09-585-858-16	Sequence 16, Appl1
33	39	36.4	1124	4	US-09-252-991A-26810	Sequence 26810, A
34	39	36.4	1235	1	US-08-680-326-36	Sequence 36, Appl1
35	38	35.5	31	1	US-08-615-770B-76	Sequence 22, Appl1
36	38	35.5	31	1	US-08-488-252-15	Sequence 15, Appl1
37	38	35.5	102	2	US-08-651-940-3	Sequence 3, Appl1
38	38	35.5	102	4	US-09-295-029-3	Sequence 3, Appl1
39	38	35.5	123	3	US-09-357-251-22	Sequence 22, Appl1
40	38	35.5	133	4	US-09-252-991A-22809	Sequence 22809, A
41	38	35.5	155	2	US-08-651-940-4	Sequence 4, Appl1
42	38	35.5	155	4	US-09-295-029-4	Sequence 4, Appl1
43	38	35.5	256	2	US-08-719-758-2	Sequence 2, Appl1
44	38	35.5	256	4	US-09-119-827-2	Sequence 2, Appl1
45	38	35.5	341	4	US-09-198-452A-293	Sequence 293, App

```

RESULT 1
5516630-2
Patent No. 5516630
APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
BAROUDY, BAHIGE M.
TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/788,262
FILING DATE: 06-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 256,135
FILING DATE: 06-OCT-1988
APPLICATION NUMBER: 654,942
FILING DATE: 27-SEP-1984
APPLICATION NUMBER: 537,911
FILING DATE: 30-SEP-1983
SEQ ID NO: 2
LENGTH: 1091
5516630-2

Query Match      100.0%; Score 107; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. NO. 7.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHIECRKPKELRLEVGKOR 20
DB      1029 SHIECRKPKELRLEVGKOR 1048

RESULT 2
US-08-475-886-2
Sequence 2, Appl1
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262052
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1

```

SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 3
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 4
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 5
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 6
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGR 20
DB 792 SHIECRKPYKELRLEVGR 811

RESULT 7
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
GENERAL INFORMATION:

APPLICANT: RAVCHANDHURI, GOPA;
EMERSON, SUZANNE, U.;
PURCELL, ROBERT, H.;
TITLE OF INVENTION: SIMIAN-HUMAN HAV
HAVING A CHIMERIC 2C PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feiler
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-42290S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGR 20
DB 792 SHIECRKPYKELRLEVGR 811

RESULT 8
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
EMERSON, SUZANNE U
PURCELL, ROBERT H
D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499

CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGR 20
DB 792 SHIECRKPYKELRLEVGR 811

RESULT 9
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
EMERSON, SUZANNE U
PURCELL, ROBERT H
D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGR 20
DB 792 SHIECRKPYKELRLEVGR 811

RESULT 10
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
EMERSON, SUZANNE U
PURCELL, ROBERT H
D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent Ver. 2.1
SEQ ID NO: 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKOR 20
|||||
792 SHIECRPKYKELRLEVGKOR 811

RESULT 11

US-08-087-016-2
Sequence 2, Application US/08087016
Patent No. 5430135
GENERAL INFORMATION:
APPLICANT: NAINAN, OMANA V.
APPLICANT: MARGOLIS, HAROLD S.
APPLICANT: ROBERTSON, BETTY H.
APPLICANT: BRINTON, MARCO H.
APPLICANT: EBERT, JAMES W.
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,016
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,828
FILING DATE: 03-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-087-016-2

Query Match 79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 3.8e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SHIECRPKYKELRLEVGKOR 20
|||||
Db 792 SHIECRPKYKELRLEVGKOR 810

RESULT 12

US-09-150-460B-11
Sequence 11, Application US/09150460B
Patent No. 6190882
GENERAL INFORMATION:
APPLICANT: Lee, Cheng-Chi
APPLICANT: Albrecht, Urs
APPLICANT: Elchele, Gregor
APPLICANT: Sun, Zhong Sheng
TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene
FILE REFERENCE: D6039
CURRENT APPLICATION NUMBER: US/09/150,460B
CURRENT FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/058,256
PRIOR FILING DATE: 1997-09-09
NUMBER OF SEQ ID NOS: 21
SEQ ID NO: 11
LENGTH: 1127
TYPE: PRT
ORGANISM: D. melanogaster Period
FEATURE:
OTHER INFORMATION: Predicted protein sequence compared with RIGU1 4.7
US-09-150-460B-11

Query Match 40.2%; Score 43; DB 3; Length 1127;
Best Local Similarity 77.8%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 ECRPKYKEL 12
|||
Db 392 ECRPKYKEL 400

RESULT 13

US-08-718-388-9
Sequence 9, Application US/08718388
Patent No. 6271362
GENERAL INFORMATION:
APPLICANT: MORIKAWA, MINORU
APPLICANT: HARADA, NAOKI
TITLE OF INVENTION: GENE ENCODING IgG Fc REGION-BINDING
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
STREET: PO BOX 747
CITY: FALLS CHURCH
STATE: VA
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/718,388
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: MURPHY JR, GERALD M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 0230-111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
INFORMATION FOR SEQ ID NO: 9:

```

; SEQUENCE CHARACTERISTICS:
;     LENGTH: 5405 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-718-388-9

```

Query Match	39.7%	Score 42.5;	DB 3;	Length 5405;
Best Local Similarity	39.1%	Pred. No. 4.4e+02;		
Matches	9;	Conservative	5;	Mismatches 4;
			Indels	5;
			Gaps	1;

Qy	2	HIECRK----	PKYELRLVGGKQ	19
		:: :	::	
Db	4824	HTVCGGAGCGPHEECRL	EDGVQ	4846

RESULT 14
 US-09-252-991A-32415
 : Sequence 32415, Application US/09252991A
 : Patent No. 6551795
 :
 : GENERAL INFORMATION:
 : APPLICANT: Marc J. Rubenfield et al.
 : TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : FILE REFERENCE: 107196.136
 : CURRENT APPLICATION NUMBER: US/09/252,991A
 : CURRENT FILING DATE: 1999-02-18
 : PRIOR APPLICATION NUMBER: US 60/074,788
 : PRIOR FILING DATE: 1998-02-18
 : PRIOR APPLICATION NUMBER: US 60/094,190
 : PRIOR FILING DATE: 1998-07-27
 : NUMBER OF SEQ ID NOS: 33142
 : SEQ ID NO 32415
 : LENGTH: 444
 : TYPE: PRT
 : ORGANISM: Pseudomonas aeruginosa
 : US-09-252-991A-32415

Query Match	39.38;	Score 42;	DB 4;	Length 444;
Best Local Similarity	50.08;	Pred. No. 34;		
Matches	7;	Conservative	4;	Mismatches 3;
				Indels 0;
				Gaps 0;

```
OY      3 IECRKPYKELREY 16
          ::|||::|||
Db      135 VQCRLPHRALRLGV 148
```

RESULT 15
 -09-328-352-6296
 Sequence 6296, Application US/09328352
 Patent No. 6562958
 GENERAL INFORMATION:
 APPLICANT: Gary L. Breton et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC99-03PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 NUMBER OF SEQ ID NOS: 8252
 SEQ ID NO 6296
 LENGTH: 472
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-352-6296

Query Match 39.3%; Score 42; DB 4; Length 472;
Best Local Similarity 40.9%; Pred. No. 37;
Matches 9; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

```
OY      1 SHIECRKPYK--ELRLEVGRQR 20  
        ||:||::|: :||:| |:  
Db      211 SHLENGRLYGQVDLKLQVAKQQ 232
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993-2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 : Search time 24.1176 seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-39
Perfect score: 107
Sequence: 1 SHIECRKPKYKELRLEVGKOR 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	107	100.0	836	US-10-272-459-40	Sequence 40, Appl
3	107	100.0	980	US-10-272-459-41	Sequence 41, Appl
4	107	100.0	2227	US-09-929-955-12	Sequence 12, Appl
5	107	100.0	2227	US-10-104-966-12	Sequence 12, Appl
6	107	100.0	2227	US-10-135-988-2	Sequence 2, Appl
7	107	100.0	2227	US-10-135-988-4	Sequence 4, Appl
8	107	100.0	2227	US-10-135-988-6	Sequence 6, Appl
9	45.5	40.2	178	US-10-091-007-124	Sequence 124, App
10	43	40.2	45	US-09-728-912-13	Sequence 13, Appl
11	43	40.2	52	US-09-728-912-12	Sequence 12, Appl
12	43	40.2	413	US-09-728-912-2	Sequence 2, Appl
13	42.5	39.7	5405	US-09-922-217-1116	Sequence 1116, Ap
14	42.5	39.7	5405	US-10-025-380-1116	Sequence 1116, Ap
15	42	39.3	1289	US-09-933-145-11	Sequence 11, Appl

16	41	38.3	22	US-10-173-461-21	Sequence 21, Appl
17	41	38.3	39	US-10-173-461-65	Sequence 65, Appl
18	41	38.3	71	US-09-864-761-43211	Sequence 43211, A
19	41	38.3	80	US-09-804-156-38	Sequence 38, Appl
20	41	38.3	80	US-10-319-519-38	Sequence 38, Appl
21	41	38.3	80	US-10-067-761-38	Sequence 38, Appl
22	41	38.3	153	US-09-533-029-96	Sequence 96, Appl
23	41	38.3	153	US-09-934-455-40	Sequence 40, Appl
24	41	38.3	153	US-10-286-264-108	Sequence 108, App
25	41	38.3	153	US-10-227-536-126	Sequence 126, App
26	41	38.3	254	US-09-820-893-94	Sequence 94, Appl
27	41	38.3	264	US-10-173-461-2	Sequence 2, Appl
28	41	38.3	315	US-09-820-893-59	Sequence 59, Appl
29	41	38.3	424	US-09-820-893-96	Sequence 96, Appl
30	41	38.3	9	US-09-978-295A-7	Sequence 7, Appl
31	41	38.3	492	US-09-978-687-7	Sequence 7, Appl
32	41	38.3	492	US-09-978-132A-7	Sequence 7, Appl
33	41	38.3	492	US-09-999-832A-7	Sequence 7, Appl
34	41	38.3	492	US-09-978-189-7	Sequence 7, Appl
35	41	38.3	492	US-09-978-608A-7	Sequence 7, Appl
36	41	38.3	492	US-09-978-585A-7	Sequence 7, Appl
37	41	38.3	492	US-09-978-191A-7	Sequence 7, Appl
38	41	38.3	492	US-09-978-403A-7	Sequence 7, Appl
39	41	38.3	492	US-09-978-564A-7	Sequence 7, Appl
40	41	38.3	492	US-09-998-833A-7	Sequence 7, Appl
41	41	38.3	492	US-09-981-915A-7	Sequence 7, Appl
42	41	38.3	492	US-09-978-824-7	Sequence 7, Appl
43	41	38.3	492	US-09-918-585A-7	Sequence 7, Appl
44	41	38.3	492	US-09-978-423A-7	Sequence 7, Appl
45	41	38.3	492	US-09-978-193A-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1:
US-10-272-459-45
Sequence 45, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANUES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: PP17955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 45
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match 100.0%: Score 107; DB 15; Length 352;
Best Local Similarity 100.0%: Pred. No. 3.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 SHIECRKPKYKELRLEVGKOR 20
Db 164 SHIECRKPKYKELRLEVGKOR 183
RESULT 2
US-10-272-459-40
Sequence 40, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANUES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

RESULT 4
US-09-929-955-12
: Sequence 12, Application US/09929955
: Patent No. US20020136740A1
GENERAL INFORMATION:
APPLICANT: Matti Salberg
APPLICANT: Catharina Hullgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
PRIOR FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17

RESULT 6
US-10-135-988-2
Sequence 2, Application US/10135988
Publication NO. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HOWDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 107; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 7
US-10-135-988-4

Sequence 4, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262053
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 107; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 8

US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262053
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 107; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 9
US-10-091-007-124

Sequence 124, Application US/10091007
Publication No. US20030170782A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Le Page, Richard W F Wells, Jeremy M
APPLICANT: Hamill, Sean B
TITLE OF INVENTION: Proteins
FILE REFERENCE: PMC/P21978WO
CURRENT APPLICATION NUMBER: US/10/091,007
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: GB 9921125.2
PRIOR FILING DATE: 1999-09-07
NUMBER OF SEQ ID NOS: 276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 124
LENGTH: 178
TYPE: PRT
ORGANISM: Streptococcus agalactiae
US-10-091-007-124

Query Match 42.5%; Score 45.5; DB 12; Length 178;
Best Local Similarity 38.5%; Pred. No. 14;
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

OY 3 IECKR-----PYKELRLEVGRKOR 19
DB 103 IKCKIKMAMLRFPFPEKELRLSAGRE 128

RESULT 10
US-09-728-912-13
Sequence 13, Application US/09728912
Patent No. US2001003643A1
GENERAL INFORMATION:

APPLICANT: Holloway, James L.
TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
FILE REFERENCE: 99-94US
CURRENT APPLICATION NUMBER: US/09/728,912
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: 60/169,252
PRIOR FILING DATE: 1999-12-06
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13
LENGTH: 45
TYPE: PRT
ORGANISM: Homo sapiens
US-09-728-912-13

Query Match 40.2%; Score 43; DB 9; Length 45;
Best Local Similarity 53.3%; Pred. No. 7.9;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 CRKPYKELRLEVGRKOR 19
DB 9 CRNPTRRLRVRGVQ 23

RESULT 11
US-09-728-912-12

```
; Sequence 12, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 52
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-728-912-12
```

```
Query Match          40.2%; Score 43; DB 9; Length 52;
Best Local Similarity 53.3%; Pred. No. 9.2;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 5 CRPKYKELRLEVQKQ 19
    |||:||||:|
Db 28 CRNPTRELRYREGVQ 42
```

```
RESULT 12
US-09-728-912-2
; Sequence 2, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 413
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-728-912-2
```

```
Query Match          40.2%; Score 43; DB 9; Length 413;
Best Local Similarity 53.3%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 5 CRPKYKELRLEVQKQ 19
    |||:||||:|
Db 376 CRNPTRELRYREGVQ 390
```

```
RESULT 13
US-09-922-217-1116
; Sequence 1116, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
```

```
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-922-217-1116
```

```
Query Match          39.7%; Score 42.5; DB 9; Length 5405;
Best Local Similarity 39.1%; Pred. No. 1.6e+03;
Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;
```

```
QY 2 HIECRK----PYKELRLEVQKQ 19
    |||:||||:|
Db 4824 HVTQCGACGACGPHECRLEDGVQ 4846
```

```
RESULT 14
US-10-025-380-1116
; Sequence 1116, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jianshun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-025-380-1116
```

```
Query Match          39.7%; Score 42.5; DB 14; Length 5405;
Best Local Similarity 39.1%; Pred. No. 1.6e+03;
Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;
```

```
QY 2 HIECRK----PYKELRLEVQKQ 19
    |||:||||:|
Db 4824 HVTQCGACGACGPHECRLEDGVQ 4846
```

```
RESULT 15
US-09-932-145-11
; Sequence 11, Application US/09932145
; Patent No. US20020161191A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
```

APPLICANT: Mintier, Gabe
APPLICANT: Kinney, Gene G
APPLICANT: Ramanaathan, Chandra S
TITLE OF INVENTION: NOVEL IMIDAZOLINE RECEPTOR HOMOLOGS
FILE REFERENCE: D0020 NP
CURRENT APPLICATION NUMBER: US/09/932,145
CURRENT FILING DATE: 2001-08-17
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 1289
TYPE: PRT
ORGANISM: Drosophila melanogaster
FEATURE:
OTHER INFORMATION: DATABASE ACCESSION NUMBER: Genbank,
OTHER INFORMATION: Accession: AAF52305
US-09-932-145-11

Query Match 39.3%; Score 42; DB 10; Length 1289;
Best Local Similarity 42.1%; Pred. No. 4.2e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
1 SHIECKRPPKELRLVSKQ 19
||:|:|:|:|:|:|:
DB 442 SHLETKKIEPLRLTYGNE 460

Search completed: October 1, 2003, 10:37:49
Job time : 24.1176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

..OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds
(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	852	1 GNNYHA	genome polypotein
2	101	100.0	1358	2 A03905	genome polypotein
3	101	100.0	2227	1 GNNYHM	genome polypotein
4	101	100.0	2227	1 GNNYHR	genome polypotein
5	101	100.0	2227	1 GNNYHR	genome polypotein
6	96	95.0	2230	1 GNNYSA	genome polypotein
7	95	94.1	2227	1 GNNYHB	genome polypotein
8	84	83.2	839	1 GNNYS2	genome polypotein
9	75	74.3	341	2 S04137	genome polypotein
10	52.5	52.0	1025	2 A54718	dihydroxyrimidine
11	47	46.5	414	2 D96838	unknown protein T2
12	45	44.6	79	2 AG1062	bacteriophage gene
13	45	44.6	321	2 E86423	probable 60S ribos
14	44	43.6	340	1 WMBE57	ribonucleoside-dip
15	44	43.6	343	2 S74937	hypothetical prote
16	44	43.6	363	2 S77299	C4-dicarboxylase-b
17	44	43.6	775	2 T48957	hypothetical prote
18	43	42.6	189	2 A12534	hypothetical prote
19	43	42.6	309	2 D75008	transposase all1756
20	43	42.6	379	2 T49919	methanol dehydroge
21	43	42.6	389	2 S41748	hypothetical prote
22	42.5	42.1	105	2 B12598	heat shock protein
23	42	41.6	109	2 S69307	hypothetical prote
24	42	41.6	165	2 S02053	probable membrane
25	42	41.6	178	2 T02335	hypothetical prote
26	42	41.6	834	2 F82673	ribosomal protein
27	41.5	41.1	304	2 H82984	hypothetical prote
28	41	40.6	83	2 T17809	hypothetical prote
29	41	40.6	159	2 C72210	conserved hypothet

30	41	40.6	213	2 H86638	hypothetical prote
31	41	40.6	229	2 T19153	hypothetical prote
32	41	40.6	237	2 C84065	transcription regu
33	41	40.6	370	2 G97001	endoglucanase fam1
34	41	40.6	561	2 G36752	unknown protein F2
35	41	40.6	622	2 T37257	hypothetical prote
36	41	40.6	897	2 T06540	axonal glycoprotein
37	41	40.6	1040	2 A34695	axonal glycoprotein
38	41	40.6	1234	2 T30254	junction protein -
39	40.5	40.1	924	2 T25007	hypothetical prote
40	40.5	40.1	932	2 T25008	hypothetical prote
41	40	39.6	177	2 C64133	mercuric resistanc
42	40	39.6	238	2 C83240	probable transcrip
43	40	39.6	259	2 F64532	conserved hypothet
44	40	39.6	272	2 T49070	probable heat choc
45	40	39.6	292	2 S23239	hypothetical prote

ALIGNMENTS

RESULT 1

GNNYHA

genome polypotein - human hepatitis A virus (strain C8326) (fragment)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03904

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra,

J. Virol. 54, 247-255, 1985

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MUID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:q329592; PIDN:AAA45470.1; PID:q329593

C:Superfamily: hepatitis A virus genome polypotein

C:Keywords: coat protein; core protein; polypotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 101; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. NO. 1.2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 2

A03905 genome polypotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Malzel Jr., J.V.; Purcell, R.H.; Feins

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polypotein

C:Keywords: coat protein; core protein; polypotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

1 Gen. Virol. 72, 1677-1683, 1991
 A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
 A:Reference number: J01080; MUID:91311420; PMID:1649901
 A:Contents: annotation
 A:Note: neither amino acid nor nucleotide sequence is given
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik
 submitted to the EMBL Data Library, May 1989
 A:Reference number: S04885
 A:Accession: S04885
 A:Molecule type: genomic RNA
 A:Residues: 1750-2164 <BAL1>
 A:Cross-references: EMBL:X15461; NID:961971; PIDN:CAA33490.1; PID:g930268
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik
 FBS Lett. 247, 425-428, 1989
 A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
 A:Reference number: S03965; MUID:89232168; PMID:2541023
 A:Accession: S03965
 A:Molecule type: genomic RNA
 A:Residues: 1960-2164 <BAL2>
 A:Cross-references: EMBL:X15461
 A:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; polyprotein
 F:1-27/Product: coat protein 1A #status predicted <C1A>
 F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-493/Product: coat protein 1C #status predicted <C1C>
 F:496-793/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: coat protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 95.0%; Score 96; DB 1; Length 2230;
 Best Local Similarity 90.0%; Pred. No. 2.5e-07;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKISLF 20
 :|||||:|||||:|||||
 DB 827 ILPPRRKMKGLFSQAKISLF 846

RESULT 7
 GNNYHB
 genome polyprotein - human hepatitis A virus (strain MBB)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
 Vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D
 A:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
 C:Accession: J50303
 R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhardt,
 Virus Res. 8, 153-171, 1987
 A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
 A:Reference number: J50303; MUID:88045071; PMID:2823500
 A:Accession: J50303
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <PAU>
 A:Cross-references: EMBL:M20273
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydr
 F:1-23/Product: coat protein 1A #status predicted <VP4>
 F:24-246/Product: coat protein 1B #status predicted <VP3>
 F:247-491/Product: coat protein 1C #status predicted <VP2>
 F:492-836/Product: coat protein 1D #status predicted <VP1>
 F:837-980/Product: core protein 2A #status predicted <P2A>
 F:981-1108/Product: core protein 2B #status predicted <P2B>
 F:1109-1438/Product: core protein 2C #status predicted <P2C>
 F:1439-1496/Product: protein 3A #status predicted <P3A>
 F:1497-1519/Product: cysteine-linked protein Vpg #status predicted <VP5>
 F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
 F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 94.1%; Score 95; DB 1; Length 2227;
 Best Local Similarity 95.0%; Pred. No. 3.6e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKISLF 20
 :|||||:|||||:|||||
 DB 823 VLPPRRKMKGLFSQAKISLF 842

RESULT 8
 GNNYB2
 genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
 N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat prot
 C:Species: simian hepatitis A virus
 A:Note: host Macaca fascicularis (cynomolgus macaque)
 C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
 C:Accession: J01180
 R:Naiman, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
 J. Gen. Virol. 72, 1685-1689, 1991
 A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus
 A:Reference number: J01180; MUID:91311421; PMID:1649902
 A:Accession: J01180
 A:Molecule type: genomic RNA
 A:Residues: 1-839 <NA1>
 A:Cross-references: GB:M59286; NID:9329599; PIDN:AAA5473.1; PID:g555083
 A:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; core protein; glycoprotein; polyprotein
 F:1-23/Product: coat protein 1A #status predicted <VP0>
 F:24-245/Product: coat protein 1B #status predicted <VP3>
 F:246-491/Product: coat protein 1C #status predicted <VP1>
 F:492-839/Product: core protein 2A (fragment) #status predicted <P2>
 F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.2%; Score 84; DB 1; Length 839;
 Best Local Similarity 83.3%; Pred. No. 8.6e-06;
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKIS 18
 :|||||:|||||:|||||
 DB 822 ILPPRRKMKGLFSQAKIS 839

RESULT 9
 S04137
 genome polyprotein - human hepatitis A virus (strain LCD-1) (fragment)
 C:Species: human hepatitis A virus
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
 C:Accession: S04137
 R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
 Nucleic Acids Res. 17, 3594, 1989
 A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir
 A:Reference number: S04137; MUID:89263805; PMID:2542903
 A:Accession: S04137
 A:Molecule type: mRNA
 A:Residues: 1-341 <AND>
 A:Cross-references: EMBL:X14666; NID:962301; PIDN:CAA32794.1; PID:g4377576
 C:Genetics:
 A:Gene: VP1
 C:Superfamily: hepatitis A virus genome polyprotein
 C:Keywords: coat protein; polyprotein
 F:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 74.3%; Score 75; DB 2; Length 341;
 Best Local Similarity 93.3%; Pred. No. 0.0001;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQA 15
 :|||||:|||||:|||||
 DB 327 VLPPRRKMKGLFSQS 341

RESULT 10

A54718
 dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 03-Jun-2002
 C:Accession: A54718
 R:Yokota, H.; Fernandez-Salguero, P.; Furuya, H.; Lin, K.; McBride, O.W.; Podschun, B.;
 J. Biol. Chem. 269, 23192-23196, 1994
 A:Title: CDNA cloning and chromosome mapping of human dihydropyrimidine dehydrogenase,
 A:Reference number: A54718; MUID:94365020; PMID:8083224
 A:Accession: A54718
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 11025 <OK>
 A:Cross-references: GB:009178; NID:9558304; PID:9558305
 A:Note: authors translated the codon GGC for residue 748 as Asp, and TTA for residue 101
 C:Genetics:
 A:Gene: GDB:DPYD
 A:Cross-references: GDB:364102; OMIM:274270
 A:Map position: 1p22-1p22
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hd
 C:Keywords: 4Fe-4S; flavoprotein; homodimer; iron-sulfur protein; NADP; oxidoreductase
 46-1005/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
 Best Local Similarity 63.2%; Pred. No. 2;
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 2 LPPPRKMKGLFSQAKISLF 20
 ||||| : |||||
 Db 176 LPEPKMKSEAVS-AKIALF 193

RESULT 11
 D96838
 unknown protein T21F11.5 [Imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: D96838
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Hughes, B.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalla,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: D96838
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-414 <STO>
 A:Cross-references: GB:AE005172; NID:96730725; PIDN:AAF27115.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: T21F11.5
 A:Map position: 1

Query Match 46.5%; Score 47; DB 2; Length 414;
 Best Local Similarity 56.2%; Pred. No. 6.2;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPRKMKGLFSQAKIS 18
 ||||| : |||||
 Db 106 PPRMDLFRKQNVLS 121

RESULT 12
 AG1062
 bacteriophage gene regulatory protein STY4826 [Imported] - Salmonella enterica subsp. en
 C:Species: Salmonella enterica subsp. enterica serovar Typhimurium
 A:Note: this species has also been called Salmonella typhimurium
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AG1062
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church
 th, T.; Connerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farr
 S.; Moulie, S.; O'Garra, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG1062
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 179 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD06948.1; PID:916505595; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY4826

Query Match 44.6%; Score 45; DB 2; Length 79;
 Best Local Similarity 63.6%; Pred. No. 2.1;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 PPRKMKGLFS 13
 ||||| : |||||
 Db 64 PPRKMGCCS 74

RESULT 13
 E86423
 probable 60S ribosomal protein L18A - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
 C:Accession: E86423
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Hughes, B.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
 ansen, N.F.; Hulzar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalla,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E86423
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-321 <STO>
 A:Cross-references: GB:AE005172; NID:910092460; PIDN:AA12862.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 44.6%; Score 45; DB 2; Length 321;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 YLPPPRKMKGLFSQAKISLF 20
 ||||| : |||||
 Db 301 YRPPTRLKTKTFRANKPNLF 320

RESULT 14
 WMB57
 ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human herpesvirus 1
 N:Alternate names: ribonucleotide reductase small chain
 C:Species: human herpesvirus 1
 C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
 R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; P
 J. Gen. Virol. 69, 1531-1574, 1988
 A:Title: The complete DNA sequence of the long unique region in the genome of herpes
 A:Reference number: A30083; MUID:88274327; PMID:2839594
 A:Accession: D30088
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A:Residues: 1-340 <MCG>
 A:Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32303.1; PID:g59540; GB:PD00317
 C:Genetics:
 A:Gene: ULA0
 C:Function:
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
 C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain
 C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; oxidc
 F:94,124,127,187,221,224/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #sta
 F:131/Active site: Tyr (stable tyrosyl radical) #status predicted

 Query Match 43.6%; Score 44; DB 1; Length 340;
 Best Local Similarity 61.5%; Pred. No. 16;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

 OY 3 PPRKMKGLFSOA 15
 |||::|||
 DB 242 PPDPRVGLFRQA 254

 ULA 15
 S74937
 hypothetical protein slr0688 - *Synechocystis* sp. (strain PCC 6803)
 C:Species: *Synechocystis* sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S74937
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
 S.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S74937
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-343 <KAN>
 A:Cross-references: EMBL:D90902; GB:AB001339; NID:g1652027; PIDN:BA16977.1; PID:d101771
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

 Query Match 43.6%; Score 44; DB 2; Length 343;
 Best Local Similarity 47.1%; Pred. No. 16;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 OY 3 PPRKMKGLFSQAISL 19
 |||::|||
 DB 240 PPDPRCKGLFRGAETI 256

 Search completed: October 1, 2003, 10:04:36
 Job time : 15.8235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 seconds
(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	852	1	P06442 hepatitis a
2	101	100.0	2227	1	P08617 hepatitis a
3	101	100.0	2227	1	P06441 hepatitis a
4	96	95.0	2226	1	P06441 hepatitis a
5	96	95.0	2226	1	P26580 hepatitis a
6	96	95.0	2226	1	P26581 hepatitis a
7	96	95.0	2226	1	P26582 hepatitis a
8	95	94.1	2230	1	P14553 simian hepa
9	84	83.2	2227	1	P13901 hepatitis a
10	84	83.2	839	1	P06441 hepatitis a
11	75	73.3	341	1	P06441 hepatitis a
12	52.5	52.0	1025	1	P13672 hepatitis a
13	45.5	45.0	1025	1	P12882 hepatitis a
14	45	44.6	178	1	Q28007 bos tauris
15	44	43.6	178	1	Q94353 oryza sativ
16	44	43.6	340	1	Q91ud4 aradidopsis
17	44	43.6	340	1	P10224 herpes simp
18	43	42.6	389	1	P06474 herpes simp
19	43	42.6	587	1	P35515 methanosarc
20	42	41.6	178	1	O53609 streptomyce
21	41	40.6	289	1	P51418 aradidopsis
22	41	40.6	359	1	P43688 mus musculi
23	41	40.6	1040	1	O54873 citriculus
24	41	40.6	1234	1	P22063 rattus norv
25	40	39.6	135	1	O62315 mus musculi
26	40	39.6	292	1	P45777 caenophiliu
27	40	39.6	310	1	P30647 haemophilus
28	40	39.6	312	1	P31330 mus musculi
29	40	39.6	353	1	Q12904 homo sapien
30	40	39.6	354	1	Q9h224 homo sapien
31	40	39.6	371	1	O9egm3 mus musculi
32	40	39.6	372	1	P43698 canis famli
33	40	39.6	372	1	P50220 homo sapien
					P23441 rattus norv

34	40	39.6	448	1	DHE4_HELPY	P55990 helicobacte
35	40	39.6	523	1	YMR9_YEAST	Q05040 saccharomyc
36	40	39.6	620	1	ARR8_SCHPO	Q9us07 schizosacch
37	40	39.6	1048	1	SR44_RAT	O63627 rattus norv
38	40	39.6	1157	1	SR44_HUMAN	O95104 homo sapien
39	39.5	39.1	1039	1	M2C1_MOUSE	O95189 mus musculi
40	39.5	39.1	1040	1	M2C1_RAT	P21139 rattus norv
41	39.5	38.6	178	1	RLIX_CASSA	O9atf5 castanea sa
42	39	38.6	204	1	RS4_TREPA	O83328 treponema p
43	39	38.6	347	1	E13A_SOYBN	O03773 glycine max
44	39	38.6	427	1	TRB1_AERPE	O9y8t5 aeropyrum p
45	39	38.6	462	1	WDR8_MOUSE	O9jmg8 mus musculi

ALIGNMENTS

RESULT 1	ID	POLG_HPVC	STANDARD;	PRT;	852 AA.
AC	P06442	083741; 083742;			
DT	01-JAN-1988	(rel. 06, Created)			
DT	01-JAN-1988	(rel. 06, Last sequence update)			
DT	16-OCT-2001	(rel. 40, Last annotation update)			
DE	Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).				
OS	Hepatitis A virus (strain CR326).				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OX	Hepatovirus.				
RN	NCBI_TaxID=12097;				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=85185648; PubMed=2985793;				
RL	Lim Meyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,				
RT	Young A., Mitra S.W.,				
RT	"Molecular cloning and partial sequencing of hepatitis A viral cDNA."				
RT	J. Virol. 54:247-255(1985).				
CC	-1 SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,				
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,				
CC	VP3, AND VP4.				
CC	-1 PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.				
CC	-1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sdb.ch/announce/				
CC	or send an email to license@sdb.ch).				
DR	EMBL; M10033; AAA45470.1; -				
DR	PIR; A03904; GNNYHA.				
KW	Polypeptide; Coat protein; Core protein.				
FT	CHAIN 1 23				
FT	CHAIN 24 245				
FT	CHAIN 246 491				
FT	CHAIN 492 836				
FT	CHAIN 837 >852				
FT	NON_TER 852 852				
FT	SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;				
Query Match	100.0%; Score 101; DB 1; Length 852;				
Best Local Similarity	100.0%; Pred. No. 9.9e-09;				
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Oy	1 VLPPPRKMKGLFSQAKISLF 20				
Db	823 VLPPPRKMKGLFSQAKISLF 842				
RESULT 2					
POLG_HPVC					

ID POLG_HPAVL STANDARD; PRT; 2227 AA.
AC P06617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [containing: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Wild type;
RX MEDLINE=97061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses.";
RT J. Virol. 61:50-59(1987).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M., Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus.";
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
[3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=8516289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Matzel J.V. Jr., Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase.";
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS. MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; M14114; AAA45475.1; -;
DR EMBL; M14707; AAA45465.1; -;
DR EMBL; M14707; AAA45466.1; ALT-INT.
DR EMBL; M16632; AAA45471.1; -;
DR PIR; A25981; GNNYTM.
DR PIR; A94149; GNNYTM.
DR PDB; 1HAV; 23-DEC-96.
DR MEROPS; C03.005; -;
DR InterPro; IPR004004; Calic1_pol_hel.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSV1r.
DR Pfam; PF00660; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KM Polypeptide; Coat protein; Core protein; Transferase; Hydrolyase; Thiol protease; 3D-structure.
KW RNA-directed RNA polymerase;
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;
Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 3
POLG_HPAVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [containing: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 1A).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., van Nest G., Dina D.;
RT "Primary structure and gene organization of human hepatitis A virus.";
RT Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS. MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
DR EMBL; M14114; AAA45475.1; -;
DR EMBL; M14707; AAA45465.1; -;
DR EMBL; M14707; AAA45466.1; ALT-INT.
DR EMBL; M16632; AAA45471.1; -;
DR PIR; A25981; GNNYTM.
DR PIR; A94149; GNNYTM.
DR PDB; 1HAV; 23-DEC-96.
DR MEROPS; C03.005; -;
DR InterPro; IPR004004; Calic1_pol_hel.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSV1r.
DR Pfam; PF00660; RNA_dep_RNA_pol; 1.

```

CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: K02990; AAA5472.1; -.
DR PIR: A03903; GNNYR.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR Polyprotein: Coat protein: Core protein: Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 837 980 CORE PROTEIN P2A.
FT CHAIN 981 1076 CORE PROTEIN P2B.
FT CHAIN 1077 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1484 PROBABLE PROTEIN P3A.
FT CHAIN 1485 1507 PROBABLE PROTEIN P3B.
FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 4
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 24a).
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12094;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91162758; PubMed-1705995;
RX Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean S.T., Jansen R.W.;
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
OC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
DR EMBL: M59810; AAA5468.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR Polyprotein: Coat protein: Core protein: Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1432 CORE PROTEIN P2C.
FT CHAIN 1433 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5
POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 43c).
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12095;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91162758; PubMed-1705995;
RX Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean S.T., Jansen R.W.;
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
OC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M59809; AAA45469.1; -
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVL.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRMKGLFSQAKISLF 20
DB 823 VLPPRRMKGLFSQAKISLF 842

ID POLG_HPAV8 STANDARD; PRT: 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_Taxid=12096;
[1]
SEQUENCE FROM N.A.
MEDLINE-91162756; PubMed-1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Crommons T., Jansen R.W.;
RT "Antigenic and genetic variation in cyclopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M59808; AAA45467.1; -
DR PDB; 1QAT; 15-MAY-00.
DR MEROPS; C03.005; -
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVL.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D68 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRMKGLFSQAKISLF 20
DB 823 VLPPRRMKGLFSQAKISLF 842

ID POLG_HPAV8 STANDARD; PRT: 2230 AA.
AC P14533;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_Taxid=12102;
[1]
SEQUENCE FROM N.A.
MEDLINE-91311420; PubMed-1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains."
RT J. Gen. Virol. 72:1677-1683(1991).
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses."
RT FEBS Lett. 247:425-428(1989).
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00924; BAA00766.1; -
DR EMBL: X15461; CAA33490.1; -
DR PIR: A30470; GNNYSA.
DR MEROPS: C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICIVIRUS.
DR Polypeptide: Coat protein; Transferase;
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249 COAT PROTEIN VP4 (P1A).
FT CHAIN 250 495 COAT PROTEIN VP2 (P1B).
FT CHAIN 496 795 COAT PROTEIN VP3 (P1C).
FT CHAIN 796 984 COAT PROTEIN VP1 (P1D).
FT CHAIN 985 1091 CORE PROTEIN P2A.
FT CHAIN 1092 1426 CORE PROTEIN P2B.
FT CHAIN 1427 1498 CORE PROTEIN P2C.
FT CHAIN 1499 1521 PROBABLE PROTEIN P3A.
FT CHAIN 1522 1741 PROBABLE PROTEIN P3B.
FT CHAIN 1742 2230 PROBABLE PROTEIN P3C.
FT CHAIN 2230 251296 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2230;
Best Local Similarity 90.0%; Pred. No. 1.8e-07;
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKKKGLFSQAKISLF 20
DB 827 ILPPRRKKKGLFSQAKISLF 846

RESULT 8
POLG_HPAVM STANDARD; PRT: 2227 AA.
AC P31901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC 081090; Q81091; Q81092; Q81093;
01-JAN-1990 (Rel. 13; Created)
01-JAN-1990 (Rel. 13; Last sequence update)
28-FEB-2003 (Rel. 41; Last annotation update)
DE Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)).
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12100;
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wiesel T., Klein R., Wimmer E.,
RA Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MBB).";
RL Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT

CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M20273; AAA45474.1; -
DR MEROPS: C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR Polypeptide: Coat protein; Core protein; Transferase;
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).
FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).
FT CHAIN 492 836 COAT PROTEIN VP3 (P1C).
FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).
FT CHAIN 981 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1496 CORE PROTEIN P2C.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3A.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3B.
FT CHAIN 1739 2227 PROBABLE PROTEIN P3C.
FT CHAIN 2227 251425 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 94.1%; Score 95; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 2.6e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRKKKGLFSQAKISLF 20
DB 823 VLPPRRKKKGLFSQAKISLF 842

RESULT 9
POLG_HPAVM STANDARD; PRT: 839 AA.
AC P31788;
AC 01-JUL-1993 (Rel. 26; Created)
01-JUL-1993 (Rel. 26; Last sequence update)
16-OCT-2001 (Rel. 40; Last annotation update)
DE Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A) (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=31707;
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nalain O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RA "Sequence analysis of a new hepatitis A virus naturally infecting
RA cynomolgus macaques (Macaca fascicularis).";
RT J. Gen. Virol. 72:1685-1689(1991).
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M59286; AAA45473.1; -
CC PIR: J01180; GNNYS2.

KM Polypeptide: Coat protein; Core protein.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 7
FT CHAIN 7 839
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACCB1E192DEC CRC64;

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 6.1e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGIFSOAKIS 18
:|||||:|||||:
822 ILPPPRKMGIFSOAKIS 839

RESULT 10
POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP3; Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LCD-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
OX NCBI_Taxid=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89263805; PubMed-2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus (HAV)."
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X14666; CA33794.1; -
CC PIR: S04137; S04137.
KM Polypeptide: Coat protein; Core protein.
FT CHAIN 1 1
FT CHAIN 2 340
FT CHAIN 341 >341
FT CHAIN 341 341
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;

Query Match 74.3%; Score 75; DB 1; Length 341;
Best Local Similarity 93.3%; Pred. No. 7.1e-05;
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGIFSOA 15
:|||||:|||||:
327 VLPPPRKMGIFSOA 341

RESULT 11
DPYD_HUMAN STANDARD; PRT; 1025 AA.
ID DPYD_HUMAN
AC Q12882; Q16694; Q16761;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Dihydropyrimidine dehydrogenase [NADP+] precursor (EC 1.3.1.2) (DPD)
DE (dihyd) (dihydrouacil dehydrogenase) (dihydrouacil dehydrogenase).
GN DPYD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-94365020; PubMed-8083224;
RA Yokota H., Fernandez-Salguero P., Furuya H., Lin K., McBride O.W.,
RA Podschun B., Schnackerz K.D., Gonzalez F.J.;
RT "CDNA cloning and chromosome mapping of human dihydropyrimidine
RT dehydrogenase, an enzyme associated with 5-fluorouracil toxicity and
RT congenital thymine uraciluria."
RL J. Biol. Chem. 269:23192-23196(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97280676; PubMed-9135003;
RA Johnson M.R., Wang K., Tillmanns S., Albin N., Diasio R.B.;
RT "Structural organization of the human dihydropyrimidine dehydrogenase
RT gene."
RL Cancer Res. 57:1660-1663(1997).
RN [3]
RP SEQUENCE OF 581-635 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE-97047101; PubMed-8892022;
RA Vreken P., van Kullenburg A.B.P., Meisema R., Smit G.P.A.,
RA Bakker H.D., de Abreu R.A., van Gennip A.H.;
RT "A point mutation in an invariant splice donor site leads to exon
RT skipping in two unrelated Dutch patients with dihydropyrimidine
RT dehydrogenase deficiency."
RL J. Inherit. Metab. Dis. 19:645-654(1996).
RN [4]
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE-92381021; PubMed-1512248;
RA Lu Z.-H., Zhang R., Diasio R.B.;
RT "Purification and characterization of dihydropyrimidine dehydrogenase
RT from human liver."
RL J. Biol. Chem. 267:17102-17109(1992).
RN [5]
RP VARIANTS ARG-29; TRP-235 AND HIS-886.
RX MEDLINE-98102836; PubMed-9439663;
RA Vreken P., van Kullenburg A.B.P., Meisema R., van Gennip A.H.;
RT "Dihydropyrimidine dehydrogenase (DPD) deficiency: identification and
RT expression of missense mutations C29R, R886H and R235W."
RL Hum. Genet. 101:333-338(1997).
RN [6]
RP VARIANTS ARG-29; TRP-235 AND HIS-886.
RX MEDLINE-97411371; PubMed-9266349;
RA Vreken P., van Kullenburg A.B.P., Meisema R., van Gennip A.H.;
RT "Identification of novel point mutations in the dihydropyrimidine
RT dehydrogenase gene."
RL J. Inherit. Metab. Dis. 20:335-338(1997).
CC -1- FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE
CC REDUCTION OF URACIL AND THYMINE. ALSO INVOLVED THE DEGRADATION OF
CC THE CHEMOTHERAPEUTIC DRUG 5-FLUOROURACIL.
CC -1- CATALYTIC ACTIVITY: 5,6-dihydrouacil + NADP(+) = uracil + NADPH.
CC -1- COFACTOR: TWO EACH OF FAD AND FMN. ALSO CONTAINS TWO FE2-S2
CC CLUSTERS. CONTAINS APPROXIMATELY 33 IRON ATOMS PER MOLECULE.
CC -1- PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF

URACIL AND THYMIDINE CATABOLISM AND IN THE PATHWAY LEADING TO THE
FORMATION OF BETA-ALANINE.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: FOUND IN MOST TISSUES WITH GREATEST ACTIVITY
CC FOUND IN LIVER AND PERIPHERAL BLOOD MONONUCLEAR CELLS.
CC -1- DISEASE: DEFECTS IN DPYD ARE THE CAUSE OF HEREDITARY THYMINE-
CC URACILURIA (ALSO KNOWN AS FAMILIAL PYRIMIDINEMIA), A DISEASE
CC CHARACTERIZED BY PERSISTENT URINARY EXCRETION OF EXCESSIVE AMOUNTS
CC OF URACIL, THYMINE AND 5-HYDROXYMETHYLURACIL. PATIENTS SUFFERING
CC FROM THIS DISEASE SHOW A SEVERE REACTION TO THE ANTICANCER DRUG 5-
CC FLUOROURACIL. THIS REACTION INCLUDES STOMATITIS, LEUKOPENIA,
CC THROMBOCYTOPENIA, HAIR LOSS, DIARRHEA, FEVER, MARKED WEIGHT LOSS,
CC CEREBELLAR ATAXIA, AND NEUROLOGIC SYMPTOMS, PROGRESSING TO
CC SEMICOMA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: U09178; AA57474.1; -
CC EMBL: U20938; AA51366.1; -
CC EMBL: X95670; CAA64973.1; -
CC PIR: A54718; A54718.
CC Genew: HGNC:3012; DPYD.
CC MIM: 274270.
CC GO: GO:0005737; C:cytoplasm; NAS.
CC GO: GO:0017113; F:5-dihydropyrimidine dehydrogenase activity; NAS.
CC GO: GO:0006214; P:thymidine catabolism; NAS.
CC GO: GO:0006212; P:uracil catabolism; NAS.
CC InterPro: IPR001450; 4fe4s_ferredoxin.
CC InterPro: IPR000759; Adrndx_reductase.
CC InterPro: IPR001295; DHO_dh.
CC InterPro: IPR005720; DHO_dhl.
CC InterPro: IPR001327; FAD_pyr_redox.
CC InterPro: IPR003009; FMN_enzyme.
CC InterPro: IPR00103; Pyridine_redox_2.
CC Pfam: PF01180; Dihydroase: 1.
CC Pfam: PF00037; fer4; 2.
CC PRINTS: PR00419; ADKRDYASE.
CC PRINTS: PR00368; FADPDR.
CC PRINTS: PR00469; PNDRTASEII.
CC TIGRFAMs: TIGR01037; PYRD_subl_fam: 1.
CC PROSITE: PS00198; 4fe4s_FERREDOXIN; 1.
CC Oxidoreductase: NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4fe-4s;
CC Disease mutation: Polymorphism.
CC PROPEP 1 3
CC FT CHAIN 4 1025 DIHYDROPYRIMIDINE DEHYDROGENASE [NADP+].
CC FT NP_BIND 335 351 NADP (POTENTIAL).
CC FT NP_BIND 471 481 FAD (POTENTIAL).
CC FT DOMAIN 661 678 URACIL-BINDING (POTENTIAL).
CC FT METAL 933 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
CC FT METAL 966 966 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
CC FT VARIANT 29 29 C -> R (In allele DPYD*9A and allele
CC DPYD*9B: loss of activity;
CC dbsnp:1801265).
CC FT VARIANT 235 235 R -> W (IN ALLELE DPYD*8: LOSS OF
CC ACTIVITY; dbsnp:1801266).
CC FT FTID-VAR_005174.
CC FT S -> N (IN ALLELE DPYD*4: LOW ACTIVITY;
CC dbsnp:1801158).
CC FT FTID-VAR_005175.
CC

FT VARIANT 543 543 I -> V (IN ALLELE DPYD*5; dbsnp:1801159).
FT FTID-VAR_005176.
FT VARIANT 732 732 V -> I (IN dBSNP:1801160).
FT FTID-VAR_014760.
FT VARIANT 886 886 R -> H (IN ALLELE DPYD*9B: 25% OF
FT ACTIVITY; dbsnp:1801267).
FT FTID-VAR_005177.
FT VARIANT 995 995 V -> F (IN ALLELE DPYD*10: LOW ACTIVITY;
FT dbsnp:1801268).
FT FTID-VAR_005178.
FT CONFLICT 910 910 S -> N (IN REF. 2).
FT SEQUENCE 1025 AA; 111374 MW; 521C9430C7E69AFA CRC64;
SQ
Query Match 52.0%; Score 52.5; DB 1; Length 1025;
Best local similarity 63.2%; Pred. No. 1.1;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
QY 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPRKMKGLFSQAKISLF 193
DPYD_BOVIN STANDARD; PRT; 1025 AA.
AC Q28007; Q9TRV4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 5-dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2) (DPD)
DE (DHPDase) (Dihydrouracil dehydrogenase) (Dihydrothymine
DE dehydrogenase).
GN DPYD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97069948; PubMed=8912928;
RA Albani N., Johnson M.R., Diasio R.B.;
RT "cdna cloning of bovine liver dihydropyrimidine dehydrogenase.";
RN DNA Seq. 6:243-250(1996).
RN [2]
RP SEQUENCE OF 668-678, AND ENZYME REGULATION.
RC TISSUE=Liver;
RX MEDLINE=92041818; PubMed=1939061;
RA Porter D.J.T., Chestnut W.G., Taylor L.C.E., Merrill B.M., Spector T.;
RT "Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.";
RN J. Biol. Chem. 266:19988-19994(1991).
CC -1- FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE
CC REDUCTION OF URACIL AND THYMINE.
CC -1- CATALYTIC ACTIVITY: 5,6-dihydrouracil + NADP(+) = uracil + NADPH.
CC -1- COFACTOR: FAD AND FMN. ALSO CONTAINS TWO 4FE-4S CLUSTERS.
CC -1- ENZYME REGULATION: INACTIVATED BY 5-iodouracil.
CC -1- PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF
CC URACIL AND THYMIDINE CATABOLISM AND IN THE PATHWAY LEADING TO THE
CC FORMATION OF BETA-ALANINE.
CC -1- SUBUNIT: Homodimer (By similarity).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL: U20981; AA840985.1; -
CC InterPro: IPR001450; 4fe4s_ferredoxin.
CC InterPro: IPR000759; Adrndx_reductase.
CC

DR InterPro: IPR001295; DHO_dh.
 DR InterPro: IPR005720; DHO_dhl.
 DR InterPro: IPR001327; FAD_pyr_redox.
 DR InterPro: IPR003009; FMN_enzyme.
 DR InterPro: IPR000103; Pyridine_redox_2.
 DR Pfam: PF01180; DHOase; 1.
 DR Pfam: PF00037; fer4; 2.
 DR Pfam: PF00070; Pyr_redox; 1.
 DR PRINTS: PR00368; ADXRDPTASE.
 DR PRINTS: PR00469; PNDROTASEII.
 DR TIGRFAMs: TIGR01037; PYRD_subl_fam; 1.
 DR PROSITE: PS00198; 4FE4S_FERRDOXIN; 1.
 DR OXIDOREDUCTASE; NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4Fe-4S.
 FT METAL 953 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
 SEQUENCE 1025 AA; 111696 MW; 5B55F93A06C47FAF CRC64;

Query Match 45.0%; Score 45.5; DB 1; Length 1025;
 Best Local Similarity 61.1%; Pred. No. 16;
 Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 2 LPPPRKMGKLFQAKISL 19
 176 LPPPKMPEAVS-AKIAL 192

RESULT 13
 ID RL1X_ORYSA STANDARD; PRT; 178 AA.
 AC 0943F3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ribosomal protein L18a.
 GN RPL18A.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Erihartoideae; Oryzaceae; Oryza.
 CC NCBI_TaxID=4530;
 RN [1]
 SEQUENCE FROM N.A.
 STRAIN=cv. Nipponbare;
 Sasaki T., Matsumoto T., Yamamoto K.;
 "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0046E05";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AP003237; BAB67920.1; -
 DR EMBL: AP003249; BAB89536.1; -
 DR Gramene: O943F3; -
 DR InterPro: IPR002670; Ribosomal_L18ae.
 DR Pfam: PF01775; Ribosomal_L18ae; 1.
 DR Ribosomal protein.
 SQ SEQUENCE 178 AA; 21400 MW; 1A508A4670C8D93B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;
 Best Local Similarity 45.0%; Pred. No. 3;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 VLPPPRKMGKLFQAKISLF 20
 158 VRPPTRLKTKTFKASRNPLF 177

RESULT 14
 ID RL1Y_ARATH STANDARD; PRT; 178 AA.
 AC OeLUD4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 60S ribosomal protein L18a-2.
 GN RPL18AB OR AT3G14600 OR MTE1.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 CC NCBI_TaxID=3702;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE-20277480; PubMed-10819329;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;
 "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
 RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC
 RT clones";
 RL DNA Res. 7:131-135(2000).
 RN [2]
 SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Shinzaki K., Davis R.W., Ecker J.R., Theologis A.;
 "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
 RT SFP consortium (Salk/Stanford/RGC)";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE L18AE FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AB023038; BAB02392.1; -
 DR EMBL: AY072340; AAU60048.1; -
 DR InterPro: IPR002670; Ribosomal_L18ae.
 DR Pfam: PF01775; Ribosomal_L18ae; 1.
 DR Ribosomal protein.
 SQ SEQUENCE 178 AA; 21309 MW; 28A1B900F3C4402B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;
 Best Local Similarity 45.0%; Pred. No. 3;
 Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 VLPPPRKMGKLFQAKISLF 20
 158 VRPPTRLKTKTFKASRNPLF 177

RESULT 15
 ID RL1R2_HSV11 STANDARD; PRT; 340 AA.
 AC P10224;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
 DE (Ribonucleotide reductase) (38 kDa subunit).
 GN UL40.

```

OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1."
RL J. Gen. Virol. 69:1531-1574(1988).
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O -> ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -1- COFACTOR: BINDS 2 IRON IONS.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14112; CAA32303.1; -.
DR PIR: D30088; WMBE57.
DR InterPro: IPR000358; Ribonucl_redctse.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 187 187 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 340 AA; 38019 MW; 4B4ED994BF74FD3F CRC64;

```

```

Query Match 43.6%; Score 44; DB 1; Length 340;
Best Local Similarity 61.5%; Pred. No. 8.7;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
3 PPPRKKKGLFSQA 15
||||:|||||
DB 242 PPPDRYGLFRQA 254

```

Search completed: October 1, 2003, 09:57:46
Job time : 7.58824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 Seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432a-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	94	12 Q9ENT9	Q9ent9 hepatitis a
2	101	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
3	101	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
4	101	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
5	101	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
6	101	100.0	94	12 Q9ENV3	Q9env3 hepatitis a
7	101	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
8	101	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
9	101	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
10	101	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
11	101	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
12	101	100.0	94	12 Q9ENV9	Q9env9 hepatitis a
13	101	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
14	101	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
15	101	100.0	94	12 Q9ENV6	Q9env6 hepatitis a
16	101	100.0	94	12 Q9ENV0	Q9env0 hepatitis a

17	101	100.0	94	12 Q9ENV6	Q9env6 hepatitis a
18	101	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
19	101	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
20	101	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
21	101	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
22	101	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
23	101	100.0	115	12 Q9DMR4	Q9dmr4 hepatitis a
24	101	100.0	116	12 Q92941	Q92941 hepatitis a
25	101	100.0	116	12 Q9W7X7	Q9w7x7 hepatitis a
26	101	100.0	116	12 Q9W7S7	Q9w7s7 hepatitis a
27	101	100.0	116	12 Q71977	Q71977 hepatitis a
28	101	100.0	116	12 Q8B8K6	Q8b8k6 hepatitis a
29	101	100.0	116	12 Q8B8K5	Q8b8k5 hepatitis a
30	101	100.0	116	12 Q8B8K4	Q8b8k4 hepatitis a
31	101	100.0	116	12 Q8B8K3	Q8b8k3 hepatitis a
32	101	100.0	126	12 Q98M19	Q98m19 hepatitis a
33	101	100.0	132	12 Q8V4I9	Q8v4i9 hepatitis a
34	101	100.0	132	12 Q8V4I6	Q8v4i6 hepatitis a
35	101	100.0	132	12 Q8V4J9	Q8v4j9 hepatitis a
36	101	100.0	132	12 Q8V4J5	Q8v4j5 hepatitis a
37	101	100.0	132	12 Q8V4M3	Q8v4m3 hepatitis a
38	101	100.0	132	12 Q8V4M7	Q8v4m7 hepatitis a
39	101	100.0	132	12 Q8V4M2	Q8v4m2 hepatitis a
40	101	100.0	132	12 Q8V4M4	Q8v4m4 hepatitis a
41	101	100.0	132	12 Q8V4M6	Q8v4m6 hepatitis a
42	101	100.0	132	12 Q8V4J7	Q8v4j7 hepatitis a
43	101	100.0	132	12 Q8V4L5	Q8v4l5 hepatitis a
44	101	100.0	132	12 Q8V4K3	Q8v4k3 hepatitis a
45	101	100.0	132	12 Q8V4K3	Q8v4k3 hepatitis a

ALIGNMENTS

RESULT 1
Q9ENT9 PRELIMINARY; PRT; 94 AA.
ID Q9ENT9
AC Q9ENT9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 32;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBD databases.
DR EMBL; AB038298; BAB11836.1; -;
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10807 MW; F19ACE91B8C4AFD CRC64;
QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 74 VLPPPRKMKGLFSQAKISLF 93
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
RESULT 2
Q9ENV9 PRELIMINARY; PRT; 94 AA.
ID Q9ENV9
AC Q9ENV9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 03;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038278; BAB1816.1; -
FT NON_TER 1 1
SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFD4AE8C4BAC CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 3

Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038282; BAB1820.1; -
FT NON_TER 1 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 4

Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;

OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 28;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038295; BAB1833.1; -
FT NON_TER 1 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 5

Q9ENV1 PRELIMINARY; PRT; 94 AA.
AC 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 49;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038306; BAB1844.1; -
FT NON_TER 1 1
SQ SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 6

Q9ENV3 PRELIMINARY; PRT; 94 AA.
AC 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 12n;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038284; BAB11822.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 7

AC O9ENV1 PRELIMINARY; PRT; 94 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 29;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038284; BAB11834.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 8

AC O9ENV8 PRELIMINARY; PRT; 94 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 22;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB038289; BAB11827.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 9

AC O9ENV2 PRELIMINARY; PRT; 94 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 15;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038285; BAB11823.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 10

AC O9ENV5 PRELIMINARY; PRT; 94 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 25;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038292; BAB11830.1; -.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGFSQAKISLF 20
 Db 74 VLPPPRKMGFSQAKISLF 93

RESULT 11

O9ENV1 PRELIMINARY; PRT; 94 AA.
 AC O9ENV1;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Hepatitis A virus.
 NCBI_TaxID=12092;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Nagasaki 18;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038286; BAB11824.1; -;
 FT NON_TER 1 94
 FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGFSQAKISLF 20
 Db 74 VLPPPRKMGFSQAKISLF 93

RESULT 12

O9ENV9 PRELIMINARY; PRT; 94 AA.
 AC O9ENV9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Hepatitis A virus.
 NCBI_TaxID=12092;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Nagasaki 21;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038288; BAB11826.1; -;
 FT NON_TER 1 94
 FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGFSQAKISLF 20

Db 74 VLPPPRKMGFSQAKISLF 93

RESULT 13

O9ENV7 PRELIMINARY; PRT; 94 AA.
 AC O9ENV7;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Hepatitis A virus.
 NCBI_TaxID=12092;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Nagasaki 06;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038280; BAB11818.1; -;
 FT NON_TER 1 94
 FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGFSQAKISLF 20
 Db 74 VLPPPRKMGFSQAKISLF 93

RESULT 14

O9ENV8 PRELIMINARY; PRT; 94 AA.
 AC O9ENV8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 OC Hepatitis A virus.
 NCBI_TaxID=12092;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Nagasaki 04;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038279; BAB11817.1; -;
 FT NON_TER 1 94
 FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGFSQAKISLF 20
 Db 74 VLPPPRKMGFSQAKISLF 93

RESULT 15

```

09ENV6
ID 09ENV6 PRELIMINARY: PRT: 94 AA.
AC 09ENV6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 07;
RA Ida S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL, AB038281; BAB11819.1; -.
NON_TER 1
NON_TER 1
SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

```

```

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VLPPPRKMKGLFSQAKISLF 20
   |||
DB 74 VLPPPRKMKGLFSQAKISLF 93

```

Search completed: October 1, 2003, 10:02:36
 Job time : 35.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds
(without alignments)
66.600 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101

Sequence: 1 VLPPRRKMKGLFSQAKISLP 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCRNUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6 5516630-2	Patent No. 5516630
2	101	100.0	2227	3 US-08-475-886-2	Sequence 2, Appl1
3	101	100.0	2227	3 US-08-475-886-4	Sequence 4, Appl1
4	101	100.0	2227	3 US-08-475-886-6	Sequence 6, Appl1
5	101	100.0	2227	3 US-08-397-232-2	Sequence 2, Appl1
6	101	100.0	2227	3 US-08-397-232-4	Sequence 4, Appl1
7	101	100.0	2227	3 US-09-171-387-2	Sequence 2, Appl1
8	101	100.0	2227	4 US-09-653-499-2	Sequence 2, Appl1
9	101	100.0	2227	4 US-09-653-499-4	Sequence 4, Appl1
10	101	100.0	2227	4 US-09-653-499-6	Sequence 6, Appl1
11	84	83.2	839	1 US-08-087-016-2	Sequence 2, Appl1
12	52.5	52.0	1025	2 US-08-304-309-2	Sequence 2, Appl1
13	52.5	52.0	1025	3 US-08-991-942-2	Sequence 2, Appl1
14	52.5	52.0	1025	3 US-09-138-103-2	Sequence 2, Appl1
15	52.5	52.0	1025	5 PCT-US95-04567-4	Sequence 4, Appl1
16	45.5	45.0	1025	5 PCT-US95-04567-2	Sequence 2, Appl1
17	45	44.6	23	1 US-07-800-364B-4	Sequence 4, Appl1
18	45	44.6	23	5 PCT-US91-03388-4	Sequence 4, Appl1
19	44	43.6	454	4 US-09-134-001C-4438	Sequence 4438, Ap
20	43.5	43.1	97	4 US-09-205-258-855	Sequence 855, App
21	42	41.6	345	4 US-09-252-991A-3076	Sequence 3076, A
22	42	41.6	1197	4 US-09-252-991A-3083	Sequence 3083, A
23	41.5	41.1	300	4 US-09-252-991A-23794	Sequence 23794, A
24	41	40.6	108	1 US-08-466-033-106	Sequence 106, App
25	41	40.6	108	2 US-08-464-733-106	Sequence 106, App
26	41	40.6	108	2 US-08-464-733-106	Sequence 106, App
27	41	40.6	108	2 US-08-461-361-106	Sequence 106, App

28	41	40.6	108	2 US-08-485-910-106	Sequence 106, App
29	41	40.6	108	5 PCT-US95-06266-89	Sequence 89, Appl
30	41	40.6	188	4 US-09-328-352-5612	Sequence 5612, Ap
31	41	40.6	616	1 US-08-638-911A-35	Sequence 35, Appl
32	41	40.6	853	1 US-08-638-911A-25	Sequence 25, Appl
33	41	40.6	853	1 US-08-638-911A-27	Sequence 27, Appl
34	41	40.6	853	1 US-08-638-911A-29	Sequence 29, Appl
35	41	40.6	853	1 US-08-638-911A-31	Sequence 31, Appl
36	41	40.6	981	4 US-09-252-991A-18616	Sequence 18616, A
37	41	40.6	2873	1 US-08-466-033-15	Sequence 15, Appl
38	41	40.6	2873	1 US-08-638-911A-2	Sequence 2, Appl1
39	41	40.6	2873	4 US-08-444-733-15	Sequence 15, Appl
40	41	40.6	2873	2 US-08-464-134-15	Sequence 15, Appl
41	41	40.6	2873	2 US-08-464-136-15	Sequence 15, Appl
42	41	40.6	2873	2 US-08-485-910-15	Sequence 15, Appl
43	41	40.6	2873	5 PCT-US95-06266-15	Sequence 15, Appl
44	41	40.6	2905	4 US-08-466-260A-401	Sequence 401, App
45	41	40.6	2905	4 US-08-488-446-401	Sequence 401, App

ALIGNMENTS

RESULT 1
5516630-2
Patent No. 5516630
APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE, STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.; BAROUDY, BAHICE M.
TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION NUMBER: US/07/788,262
FILING DATE: 06-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 256,135
FILING DATE: 06-OCT-1988
APPLICATION NUMBER: 654,942
FILING DATE: 27-SEP-1984
APPLICATION NUMBER: 537,911
FILING DATE: 30-SEP-1983
SEQ ID NO: 2
LENGTH: 1091
5516630-2

Query Match 100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 9.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQAKISLP 20
DB 1060 VLPPRRKMKGLFSQAKISLP 1079

RESULT 2
US-08-475-886-2
Sequence 2, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1

```
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLP...KMG...F...S...Q...A...K...I...S...L...F... 20
      823 VLP...KMG...F...S...Q...A...K...I...S...L...F... 842
```

```
RESULT 3
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
```

```
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLP...KMG...F...S...Q...A...K...I...S...L...F... 20
      823 VLP...KMG...F...S...Q...A...K...I...S...L...F... 842
```

```
LT 4
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLP...KMG...F...S...Q...A...K...I...S...L...F... 20
      823 VLP...KMG...F...S...Q...A...K...I...S...L...F... 842
```

```
RESULT 5
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
```

```
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLP...KMG...F...S...Q...A...K...I...S...L...F... 20
      823 VLP...KMG...F...S...Q...A...K...I...S...L...F... 842
```

```
RESULT 6
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Oy      1 VLPPPRKMG LFSQAKISLF 20
          |||||
Db      823 VLPPPRKMG LFSQAKISLF 842

```

RESULT 7
US-09-171-387-2

;; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; ADDITIONAL INFORMATION: 301

APPLICANT: RAYCHAUDHURI, GOPA;
EMERSON, SUZANNE, U.;
PROJECT: "DORRIS"

TITLE OF INVENTION: SIMIAN-HUMAN HAV
HAVING A CHIMERIC 2C PROTEIN

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
ADDRESS: 345 BAY STREET
SUITE 200
SAN FRANCISCO, CA 94111-4209
TELEPHONE: 415 774 2000
FAX: 415 774 2001
E-MAIL: MORGAN@FINNEGAN.COM

COMPUTER READABLE FORM:

```

;
; MEDIUM TYPE: FLOPPY DISK
;
; COMPUTER: IBM PC COMPATIBLE
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: MICROSOFT WORD 97
;

```

APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999

;
 ; PRIOR APPLICATION DATA:
 ;
 APPLICATION NUMBER: PCT/US97/06506
 FILING DATE: 10 DEC 1997

FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
DATE 04/19/96 09:54:00

REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229051

```

; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:

```

SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match	100.0%	Score 101; DB 3;	Length 2227;
Best Local Similarly	100.0%;	Pred. No. 2.1e-08;	
Matches 20; Conservative	0;	Mismatches 0;	Indels 0; Gaps 0

QY	1	VLPPPRKMGFLFSQAKISLF	20
Db	823	VLPPPRKMGFLFSQAKISLF	842

RESULT 8
US-09-653-499-2

```

; Sequence 2, Application US/09653499
; Patent No. 6423318
: GENERAL INFORMATION:

```

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

```

: CURRENT FILING DATE: 2000-08-31
: PRIOR APPLICATION NUMBER: 08/475,886
: PRIOR FILING DATE: 1999-08-09
: PRIOR APPLICATION NUMBER: 07/947,338
: PRIOR FILING DATE: 1992-09-18
: PRIOR APPLICATION NUMBER: 08/397,232
: PRIOR FILING DATE: 1995-03-10
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2

```

US-09-653-499-2

Query Match	100.0%	Score 101;	DB 4;	Length 2227;
Best Local Similarity	100.0%	Pred. No. 2.1e-08;		
Matches 20; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Qy      1 VLPPPRKMGFSQAKISLF 20
          |||||
Db      823 VLPPPRKMGFSQAKISLF 842

```

RESULT 9
US-09-653-499-4

```

; Sequence 4, Application US/096553499
; Patent No. 6423318
; GENERAL INFORMATION:

```

;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE D
;; APPLICANT: PURCELL, ROBERT H

```

: APPLICANT: D'HONDT, ERIC
:
: TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
:
: FILE REFERENCE: 20264262US2
:

```

; CURRENT APPLICATION NUMBER: US/09/65
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886

; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18

PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1

US-09-653-499-4

```

Best Local Similarity 100.0%; Pred. NO. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db 823 VLPPPKMKGLFSQAKISLF 842

RESULT 10
US-09-653-499-6
Sequence 6 Annotation TC/00653400

```

; sequence 8, Application 05/05
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: EINHORNSEB ANN W

```

APPLICANT: EMERSON, SUZANNE D
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'AMONTE, ERIC
APPLICANT: FUNKHOUSER, ANN W

PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO: 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||||
823 VLPPPRKMKGLFSQAKISLF 842

RESULT 11
US-08-087-016-2
Sequence 2, Application US/08087016
Patent No. 5430135
GENERAL INFORMATION:
APPLICANT: NAINAN, OMANA V.
APPLICANT: MARGOLIS, HAROLD S.
APPLICANT: ROBERTSON, BETTY H.
APPLICANT: BRINTON, MARGO H.
APPLICANT: EBERT, JAMES W.
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,016
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,828
FILING DATE: 03-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-087-016-2

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 5.5e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKIS 18
:|||||:|||||:|||||
Db 822 ILPPPRKMKGLFSQAKIS 839

RESULT 12
US-08-304-309-2
Sequence 2, Application US/08304309
Patent No. 5856454
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1025 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-304-309-2

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGLFSQAKISLF 20
||||| : |||||
Db 176 LPPPRKMKSEAVS-AKTALF 193

RESULT 13
US-08-991-942-2
Sequence 2, Application US/08991942
Patent No. 6015673
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Stewart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493


```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,942
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-942-2

Query Match          52.0%; Score 52.5; DB 3; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY      2 LPPPRKMGKGFSAQKISLF 20
DB      176 LPPPEKMESEAYS-AKIALF 193

RESULT 14
US-09-138-103-2
; Sequence 2, Application US/09138103A
; Patent No. 6232448
; GENERAL INFORMATION:
; APPLICANT: Yoshikubo, Takashi
; APPLICANT: Hasegawa, Masami
; TITLE OF INVENTION: Immunological Materials and Methods for Detecting
; FILE REFERENCE: 09/138,103 Yoshikubo, et al.
; CURRENT APPLICATION NUMBER: US/09/138,103A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 97114630.3
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-138-103-2

Query Match          52.0%; Score 52.5; DB 3; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY      2 LPPPRKMGKGFSAQKISLF 20
DB      176 LPPPEKMESEAYS-AKIALF 193

RESULT 15
PCT-US95-04567-4
; Sequence 4, Application PC/TUS9504567
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
```

```

;
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: Patentin Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04567
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,357
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UOAB025P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04567-4

Query Match          52.0%; Score 52.5; DB 5; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.4;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY      2 LPPPRKMGKGFSAQKISLF 20
DB      176 LPPPEKMESEAYS-AKIALF 193
```

Search completed: October 1, 2003, 10:06:29
Job time : 13.7059 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 : Search time 24.1176 seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

al number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	352	15 US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	15 US-10-272-459-41	Sequence 12, Appl
3	101	100.0	2227	10 US-09-928-955-12	Sequence 11, Appl
4	101	100.0	2227	14 US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	14 US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	14 US-10-135-988-4	Sequence 4, Appl
7	101	100.0	2227	14 US-10-135-988-6	Sequence 6, Appl
8	70	69.3	836	15 US-10-272-459-40	Sequence 40, Appl
9	52.5	52.0	1025	9 US-09-854-886-2	Sequence 2, Appl
10	43.5	43.1	673	16 US-10-176-306-2	Sequence 2, Appl
11	43.5	43.1	97	12 US-09-933-767-855	Sequence 855, App
12	43.5	43.1	97	15 US-10-023-282-855	Sequence 855, App
13	42	41.6	79	15 US-10-106-698-4310	Sequence 4310, Ap
14	42	41.6	449	15 US-10-156-761-13588	Sequence 13588, A
15	42	41.6	2053	12 US-10-323-450-12	Sequence 12, Appl

16	42	41.6	2053	14 US-10-017-216-2	Sequence 2, Appl
17	41	40.6	2842	12 US-10-224-999A-3479	Sequence 3479, Ap
18	41	40.6	2905	8 US-08-424-550B-401	Sequence 401, App
19	41	40.6	2910	12 US-09-828-498-2	Sequence 2, Appl
20	40	39.6	166	10 US-09-851-026-11	Sequence 11, Appl
21	40	39.6	166	11 US-09-930-169-3	Sequence 3, Appl
22	40	39.6	272	11 US-09-934-455-140	Sequence 35, Appl
23	40	39.6	312	10 US-09-851-026-35	Sequence 36, Appl
24	40	39.6	312	10 US-09-851-026-36	Sequence 2, Appl
25	40	39.6	318	10 US-09-947-971-2	Sequence 472, App
26	40	39.6	320	15 US-10-102-806-472	Sequence 76, Appl
27	40	39.6	371	9 US-09-320-337-76	Sequence 4933, Ap
28	40	39.6	666	9 US-09-815-242-4933	Sequence 10663, A
29	40	39.6	657	11 US-09-815-242-10663	Sequence 1915, Ap
30	40	39.6	871	9 US-09-991-936-1915	Sequence 32, Appl
31	39.5	39.1	184	9 US-09-140-668A-32	Sequence 2, Appl
32	39.5	39.1	390	15 US-10-283-013-2	Sequence 195, App
33	39.5	39.1	467	10 US-09-903-340-195	Sequence 195, App
34	39.5	39.1	467	10 US-09-909-088B-195	Sequence 195, App
35	39.5	39.1	467	10 US-09-905-291A-195	Sequence 195, App
36	39.5	39.1	467	10 US-09-981-876-134	Sequence 134, App
37	39.5	39.1	467	10 US-09-902-853-195	Sequence 195, App
38	39.5	39.1	467	10 US-09-907-834-195	Sequence 195, App
39	39.5	39.1	467	10 US-09-907-841-195	Sequence 195, App
40	39.5	39.1	467	11 US-09-904-011-195	Sequence 195, App
41	39.5	39.1	467	11 US-09-906-742-195	Sequence 195, App
42	39.5	39.1	467	11 US-09-148-545-134	Sequence 134, App
43	39.5	39.1	467	11 US-09-906-838-195	Sequence 195, App
44	39.5	39.1	467	11 US-09-907-613-195	Sequence 195, App
45	39.5	39.1	467	11 US-09-907-942-195	Sequence 195, App

ALIGNMENTS

RESULT 1
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHANTRES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI1955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 KDa
US-10-272-459-45

Query Match 100.0%; Score 101; DB 15; Length 352;
Best local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 195 VLPPPRKMKGLFSQAKISLF 214
RESULT 2
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHANTRES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
```

```
Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPIRMKGLFSQAKISLF 20
      |||
Db      823 VLPPIRMKGLFSQAKISLF 842
```

```
RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
```

```
Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPIRMKGLFSQAKISLF 20
      |||
Db      823 VLPPIRMKGLFSQAKISLF 842
```

```
RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPIRMKGLFSQAKISLF 20
      |||
Db      823 VLPPIRMKGLFSQAKISLF 842
```

```
RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HW-175
US-10-135-988-2
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPIRMKGLFSQAKISLF 20
      |||
Db      823 VLPPIRMKGLFSQAKISLF 842
```

```
RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQAKISLF 20
DB 823 VLPPRRKMKGLFSQAKISLF 842

RESULT 7
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN M
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VLPPRRKMKGLFSQAKISLF 20
823 VLPPRRKMKGLFSQAKISLF 842

RESULT 8
US-10-272-459-40
Sequence 40, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANTES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: PPI7955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 40
LENGTH: 836
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-40

Query Match 69.3%; Score 70; DB 15; Length 836;
Best Local Similarity 92.9%; Pred. No. 0.019;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQ 14
DB 823 VLPPRRKMKGLFSQ 836

RESULT 9
US-09-854-886-2
Sequence 2, Application US/09854886
Patent No. US20020072080A1
GENERAL INFORMATION:
APPLICANT: Yoshikubo, Takashi
APPLICANT: Hasegawa, Masami
TITLE OF INVENTION: Immunological Material and Methods for Detecting
FILE REFERENCE: 100554-32887
CURRENT APPLICATION NUMBER: US/09/854,886
CURRENT FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 09/138,103
PRIOR FILING DATE: 1998-08-21
PRIOR APPLICATION NUMBER: 97114630.3
PRIOR FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1025
TYPE: PRT
ORGANISM: Homo sapiens
US-09-854-886-2

Query Match 52.0%; Score 52.5; DB 9; Length 1025;
Best Local Similarity 63.2%; Pred. No. 11;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPEKMKGLFSQAKISLF 20
DB 176 LPPEKMKGLFSQAKISLF 193

RESULT 10
US-10-176-306-2
Sequence 2, Application US/10176306
Publication No. US20030130485A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
APPLICANT: Curtis, Rory A. J.
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Bandaru, Rajasekhar
APPLICANT: Kapeller-Libermann, Rosana
TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF
FILE REFERENCE: 10448-195001
CURRENT APPLICATION NUMBER: US/10/176,306
CURRENT FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: 10/001,137
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: PCT/US01/45291
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/248,362
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/248,331
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/248,365
PRIOR FILING DATE: 2000-11-14
PRIOR APPLICATION NUMBER: 60/250,077
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/250,327
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 60/250,176
PRIOR FILING DATE: 2000-11-30

;; PRIOR APPLICATION NUMBER: 10/023,617
;; PRIOR FILING DATE: 2001-12-18
;; PRIOR APPLICATION NUMBER: PCT/US01/49416
;; PRIOR FILING DATE: 2001-12-18
;; PRIOR APPLICATION NUMBER: 60/256,249
;; PRIOR FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: 60/256,405
;; PRIOR FILING DATE: 2000-12-18
;; PRIOR APPLICATION NUMBER: 10/083,248
;; PRIOR FILING DATE: 2001-10-22
;; PRIOR APPLICATION NUMBER: PCT/US01/46717
;; PRIOR FILING DATE: 2001-10-22
;; PRIOR APPLICATION NUMBER: 60/242,324
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/242,518
;; PRIOR FILING DATE: 2000-10-23
;; PRIOR APPLICATION NUMBER: 60/241,989
;; PRIOR FILING DATE: 2000-10-20
;; NUMBER OF SEQ ID NOS: 86
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO: 2
;; LENGTH: 673
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-176-306-2

Query Match 43.6%; Score 44; DB 16; Length 673;
Best Local Similarity 43.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 3 PPPKKMGFSQAKIS 18
||||:||||:|
Db 507 PPPGGKSFSSRELT 522

RESULT 11
US-09-933-767-855
; Sequence 855, Application US/09933767
; Publication No. US20030181692A1
; GENERAL INFORMATION:
; APPLICANT: NI et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P2
; CURRENT APPLICATION NUMBER: US/09/933,767
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: PCT/US01/05614
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,836
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/193,170
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/205,258
; PRIOR FILING DATE: 1998-12-04
; PRIOR APPLICATION NUMBER: PCT/US98/11422
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/048,885
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,375
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,881
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,880
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,896
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06

;; PRIOR APPLICATION NUMBER: 60/048,894
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,971
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,964
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,882
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,899
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,893
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,900
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,901
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,892
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,915
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,019
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,970
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,972
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,916
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,373
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,875
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/049,374
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,917
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,949
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,974
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,883
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,897
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,898
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,962
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,963
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,877
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/048,878
;; PRIOR FILING DATE: 1997-06-06
;; PRIOR APPLICATION NUMBER: 60/058,054
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,064
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/068,053
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/070,923
;; PRIOR FILING DATE: 1997-12-18
;; PRIOR APPLICATION NUMBER: 60/073,160
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/073,159
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/073,165
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/073,164
;; PRIOR FILING DATE: 1998-01-30
;; PRIOR APPLICATION NUMBER: 60/085,925
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/085,921

```

; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-933-767-855

Query Match      43.1%; Score 43.5; DB 12; Length 97;
Best Local Similarity 47.4%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

2 LPPPRKMGVLA-IRLSLF 69
52 VPPPLKMGVLA-IRLSLF 69
Db

RESULT 12
US-10-023-282-855
; Sequence 855, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-282-855

Query Match      43.1%; Score 43.5; DB 15; Length 97;
Best Local Similarity 47.4%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

2 LPPPRKMGVLA-IRLSLF 69
52 VPPPLKMGVLA-IRLSLF 69
Db

RESULT 13
US-10-106-698-4310
; Sequence 4310, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
```

```
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: US 60/157,137
;; PRIOR FILING DATE: 1999-09-29
;; PRIOR APPLICATION NUMBER: US 60/163,280
;; PRIOR FILING DATE: 1999-11-03
;; NUMBER OF SEQ ID NOS: 8564
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 4310
;; LENGTH: 79
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (13)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: MISC_FEATURE
;; LOCATION: (79)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4310
```

Query Match 41.6%; Score 42; DB 15; Length 79;

Best Local Similarity 55.6%; Pred. No. 32;

Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

OY 3 PPRKMKGLFSQAKISLF 20

DB 34 PPRKME-FRTASIRLF 49

RESULT 14

US-10-156-761-13588

;; Sequence 13588, Application US/10156761

;; Publication No. US20030119018A1

;; GENERAL INFORMATION:

;; APPLICANT: OMURA, SATOSHI

;; APPLICANT: IKEDA, HARUO

;; APPLICANT: ISHIKAWA, JUN

;; APPLICANT: HORIKAWA, HIROSHI

;; APPLICANT: SHIBA, TADAYOSHI

;; APPLICANT: SAKAKI, YOSHITOKI

;; APPLICANT: HATORI, MASAHIRA

;; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

;; FILE REFERENCE: 249-262

;; CURRENT APPLICATION NUMBER: US/10/156,761

;; PRIOR FILING DATE: 2002-05-29

;; PRIOR APPLICATION NUMBER: JP 2001-204089

;; PRIOR FILING DATE: 2001-05-30

;; PRIOR APPLICATION NUMBER: JP 2001-272697

;; PRIOR FILING DATE: 2001-08-02

;; NUMBER OF SEQ ID NOS: 15109

;; SEQ ID NO 13588

;; LENGTH: 449

;; TYPE: PRT

;; ORGANISM: Streptomyces avermitilis

US-10-156-761-13588

Query Match 41.6%; Score 42; DB 15; Length 449;

Best Local Similarity 44.4%; Pred. No. 2e+02;

Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 VLPPRRMKGLFSQAKIS 18

DB 381 LMAPRGVRLFSKAGVT 398

RESULT 15

US-10-325-430-12

;; Sequence 12, Application US/10325430

;; Publication No. US20030153525A1

;; GENERAL INFORMATION:

;; APPLICANT: Millennium Pharmaceuticals, Inc

;; APPLICANT: Silos-Santiago, Immaculada

;; APPLICANT: Rosenfeld, Julie Beth

;; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING

;; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,

;; FILE REFERENCE: MP101-294PIRNM

;; CURRENT APPLICATION NUMBER: US/10/325,430

;; CURRENT FILING DATE: 2002-12-19

;; PRIOR APPLICATION NUMBER: US 60/341,953

;; PRIOR FILING DATE: 2001-12-19

;; NUMBER OF SEQ ID NOS: 21

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 12

;; LENGTH: 2053

;; TYPE: PRT

;; ORGANISM: Homo Sapiens

US-10-325-430-12

Query Match 41.6%; Score 42; DB 12; Length 2053;

Best Local Similarity 61.5%; Pred. No. 9.8e+02;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 PPRKMKGLFSQAK 16

DB 1259 PAKKKGLFSRRK 1271

Search completed: October 1, 2003, 10:37:50

Job time : 24.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-43
Perfect score: 101
Sequence: 1 FSOAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
al number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	18 AAW42926	Immunogenic Hepati
2	101	100.0	21	22 AAB69443	Synthetic HAV p2a
3	101	100.0	854	6 AAP50287	Sequence encoded b
4	101	100.0	1077	20 AAW95559	A partial hepatitis
5	101	100.0	1091	14 AAR32426	Translated from 5'
6	101	100.0	2227	11 AAR05697	Attenuated hepatitis
7	101	100.0	2227	18 AAW34074	Hepatitis A virus
8	101	100.0	2227	21 AAB18607	Amino acid sequenc
9	101	100.0	2227	21 AAB18608	Amino acid sequenc

10	101	100.0	2227	21 AAB18609	Amino acid sequenc
11	101	100.0	2227	23 AAG31727	Wild-type Hepatitis
12	101	100.0	2227	23 AAG31728	Hepatitis A virus
13	101	100.0	2227	23 AAG31729	Attenuated Hepati
14	101	100.0	2227	23 AAE19899	Hepatitis A virus
15	101	100.0	2227	24 ABU08639	Wild type human he
16	101	100.0	2227	24 ABU08640	Attenuated (pass3
17	101	100.0	2227	24 ABU08641	Attenuated hepatitis
18	98	97.0	2227	7 AAP60066	Sequence of viral
19	97	96.0	366	6 AAP50230	Sequence of hepati
20	97	96.0	993	6 AAP50116	Sequence of hepati
21	97	96.0	993	6 AAP50231	Sequence encoded b
22	49.5	49.0	20	19 AAW76539	Plasmid pGEM2-HAV-
23	47	46.5	518	20 AAY27340	Group B Streptococ
24	47	46.5	816	20 AAY27339	Streptococcus poly
25	47	46.5	846	23 ABP30554	Streptococcus poly
26	47	46.5	847	23 ABP27030	Streptococcus poly
27	47	46.5	847	23 ABP29773	Streptococcus poly
28	43	42.6	171	23 ABP40360	Staphylococcus epi
29	42	41.6	20	18 AAW42925	Immunogenic Hepat
30	42	41.6	20	18 AAB69442	Synthetic HAV p2a
31	42	41.6	125	23 AAB49486	Listeria monocytrog
32	42	41.6	172	23 AAB25691	Streptococcus poly
33	42	41.6	503	22 AAG15762	Novel human diagno
34	41	40.6	144	22 AAB96238	Putative P. abysal
35	41	40.6	173	22 AAB49279	Protein encoded by
36	41	40.6	238	22 AAB68889	Drosophila melano
37	41	40.6	2431	13 AAR25138	SPV4 non-structura
38	40	39.6	105	21 AAG19719	Arabidopsis thalia
39	40	39.6	105	21 AAG61014	Arabidopsis thalia
40	40	39.6	122	21 AAG41997	Arabidopsis thalia
41	40	39.6	173	22 AAV29276	Human PRO polypept
42	40	39.6	173	22 AAB49278	Protein encoded by
43	40	39.6	173	24 ABU71364	Human PRO4423 prot
44	40	39.6	173	24 ABU65821	Human secreted/tra
45	40	39.6	173	24 ABU65154	Novel human secret

ALIGNMENTS

RESULT 1
ID AAW42926 standard; peptide: 20 AA.
XX
AC AAW42926;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1319.
XX
KW Immunogenic peptide; immunogenic epitope; p2a protein;
KW Immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
PR 19-APR-1996; 96US-0015644.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudiyakov YE;
XX
DR WPI, 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal
 PS Claim 18; Page 112; 140pp; English.
 XX
 CC Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. The present peptide
 CC is derived from amino acids 834-853, and has a reactivity of 27.1% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal.
 CC
 XX Sequence 20 AA:
 SQ
 Query Match 100.0%; Score 101; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 7.7e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 FSOAKISLFYTEHEHMKFS 20
 1 FSOAKISLFYTEHEHMKFS 20
 Db
 RESULT 2
 AAB69443
 ID AAB69443 standard; Peptide; 21 AA.
 XX AAB69443:
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 43.
 XX
 KM Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KM antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 XX
 PN WO200105824-A2.
 XX
 PD 25-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-US19267.
 XX
 PF 15-JUL-1999; 99US-0144412.
 XX
 (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields RA, Khudyakov YE;
 XX
 DR WPI; 2001-112681/12.
 XX
 PT Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines -
 XX
 PS Claim 13; Page 95; 130pp; English.
 XX
 CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC Igm antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy
 CC end of the peptides enhances the Igm antibody reactivity.
 CC
 XX Sequence 21 AA:
 SQ
 Query Match 100.0%; Score 101; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8.1e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 FSOAKISLFYTEHEHMKFS 20
 1 FSOAKISLFYTEHEHMKFS 20
 Db
 RESULT 3
 AAP50287
 ID AAP50287 standard; Protein; 854 AA.
 XX AAP50287:
 XX
 AC AAP50287;
 XX
 DT 25-MAR-2003 (updated)
 DT 30-NOV-1991 (first entry)
 XX
 DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the
 DE genome 5' terminus to the end of the area corresponding to the
 DE capsid protein region of poliovirus RNA.
 XX
 KM Hepatitis A virus assay; antigen; antibody.
 XX
 OS Hepatitis A virus.
 OS
 XX
 PN WO8501517-A.
 XX
 PD 11-APR-1985.
 XX
 PF 27-SEP-1984; 84WO-US01552.
 XX
 PF 30-SEP-1983; 83US-0537911.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Tlcehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racanelli VR;
 XX
 DR WPI; 1985-098846/16.
 DR N-PSDB: AAN50330.
 XX
 FT New hepatitis A virus cDNA - useful in assays for the virus and
 FT for prodn. of the viral antigen and antibodies to it
 XX
 PS Example; Fig 7; 60pp; English.
 XX
 CC The inventors claim HAV cDNA and a method for producing it, whereby
 CC large amts. can be obtd. economically. The cDNA is useful in the
 CC assay for detection of HAV quickly and easily and with high
 CC sensitivity and specificity. The HAV cDNA is also used in the prodn.
 CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 854 AA:
 Query Match 100.0%; Score 101; DB 6; Length 854;
 Best Local Similarity 100.0%; Pred. No. 5e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 FSOAKISLFYTEHEHMKFS 20
 1 FSOAKISLFYTEHEHMKFS 20
 Db 834 FSOAKISLFYTEHEHMKFS 853
 RESULT 4

```

AAW95559
ID AAW95559 standard; Protein; 1077 AA.
XX
AC AAW95559;
XX
DT 28-APR-1999 (first entry)
XX
DE A partial hepatitis A virus (HAV) protein.
XX
KM Hepatitis A virus protein; HAV; P2 region;
XX cell-culture-adapted HAV strain; infection; accelerated growth.
XX
OS Hepatitis A virus.
XX
PN US5849562-A.
XX
PD 15-DEC-1998.
XX
PF 06-JUN-1995; 95US-0468926.
XX
PR 06-NOV-1991; 91US-0788262.
XX 30-SEP-1983; 83US-0537911.
XX 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 06-JUN-1995; 95US-0468926.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Emerson SU, Purcell RH;
XX
DR WPI; 1999-094412/08.
XX N-PSDB; AAX01006.
XX
PT Chimeric hepatitis A virus strains - with P2 region from
XX cell-culture-adapted strain in wild-type genome
XX
PS Disclosure; Fig 7A-L; 36pp; English.
XX
CC The present sequence represents a partial hepatitis A virus (HAV)
XX protein. The specification describes a DNA construct consisting
XX of a wild-type HAV genome in which the P2 region is replaced by the
XX P2 region from a cell-culture-adapted HAV strain. The construct is
XX used to demonstrate that mutations in the P2 region of a
XX cell-culture-adapted HAV strain are sufficient for establishment of
XX infection and accelerated growth in cell culture.
XX
SQ Sequence 1077 AA;
XX
Query Match 100.0%; Score 101; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKPS 20
DB 1057 FSOAKISLFYTEHEIMKPS 1076
XX
RESULT 5
ID AAR32426 standard; Protein; 1091 AA.
XX
AC AAR32426;
XX
DT 25-MAR-2003 (updated)
XX 17-DEC-2001 (updated)
XX 10-JUN-1993 (first entry)
XX
DE Translated from 5' region of Hepatitis A Virus genomic clone.
XX
KM HAV HM-175; chronic liver disease; picornavirus.
XX
OS Hepatitis A virus.
XX

```

```

FH Key Location/Qualifiers
FT Region 238..1091
FT FT /label=ORF
FT FT /note="second putative initiation codon at
FT FT position 240"
FT FT 1..711
FT FT /note="X's correspond to nonsense codons,
FT FT i.e. this region is not an ORF"
XX
PN USN7788262-N.
XX
PD 15-DEC-1992.
XX
PF 30-SEP-1983; 83US-0536911.
XX
PR 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 30-SEP-1983; 83US-0536911.
XX 06-NOV-1991; 91US-0788262.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH, Racanelli VR;
XX Baroudy BM, Emerson SU;
XX
DR WPI; 1993-067429/08.
XX N-PSDB; AAO36934.
XX
PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
XX of antigen and antibodies
XX
PS Disclosure; Fig 7; 65pp; English.
XX
CC HAV virion RNA was extracted from the livers of marmosets which had
XX been inoculated with HAV (the HAV had previously been passaged twice
XX in marmosets). The RNA was used to prepare ds cDNA clones by
XX standard methods. Clones contg. inserts which hybridised to RNA from
XX HAV-infected African Green Monkey Kidney cells were selected for
XX further analysis. A 7.4kb restriction map (about 9% of the HAV
XX genome) was constructed from 5 overlapping inserts. The sequence of
XX the first 3.3kb (approx.) from the 5' terminus was determined. An
XX amino acid sequence was decoded from the entire clone and an open
XX reading frame was identified starting at position 238. A comparison
XX of the predicted HAV amino acid sequences with the known capsid
XX protein sequences of other picornaviruses (poliovirus, foot and
XX mouth disease virus and encephalomyelitis virus) revealed areas of
XX local homology.
XX
CC (Note: Revised entry submitted to correct the patent number format of
XX US Government-owned WPIs applications to prevent clashes with ongoing US
XX granted patent numbers. For further information please visit the Derwent
XX web site at www.derwent.com/dwpi/updates/ntis\_us.html.)
XX
SQ Sequence 1091 AA;
XX
Query Match 100.0%; Score 101; DB 14; Length 1091;
Best Local Similarity 100.0%; Pred. No. 6.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKPS 20
DB 1071 FSOAKISLFYTEHEIMKPS 1090
XX
RESULT 6
ID AAR05697 standard; Protein; 2227 AA.
XX
AC AAR05697;
XX
DT 25-MAR-2003 (updated)
XX 15-AUG-1990 (first entry)
XX

```

	DE	Attenuated hepatitis A virus.
XX	KW	Hepatitis A virus; vaccine; attenuated.
OS	OS	Hepatitis A virus, strain HM-175.
XX	FH	Key
FT	Region	Location/Qualifiers 1..23
FT	Region	/label=VP4 = 1A
FT	Region	24..245
FT	Region	/label=VP2 = 1B
FT	Region	246..491
FT	Region	/label=VP3 = 1C
FT	Region	492..791
FT	Region	/label=VP1 = ID
FT	Region	792..980
FT	Region	/label=2A
FT	Region	981..1087
FT	Region	/label=2B
FT	Region	1088..1422
FT	Region	/label=2C
FT	Region	1423..1496
FT	Region	/label=3A
FT	Region	1497..1519
FT	Region	/label=3B = VPg
FT	Region	1520..1738
FT	Region	/label=3C
FT	Region	1739..2227
FT	Region	/label=3D
PN	US4894228-A.	
PD	16-JAN-1990.	
PF	12-JUL-1988;	88US-0217824.
PR	12-JUL-1988;	88US-0217824.
PR	12-JUL-1988;	88US-0652967.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICE.	
PI	Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Felstone SM,	
PI	Daeam R J., Gust ID;	
DR	WPJ. 1990-075557/10.	
DR	N-PSTDB; AAQ03512.	
Vaccine against hepatitis A virus infection - comprises novel attenuated hepatitis A virus strain.		
Claim 1; Fig 1; 18pp; English.		
The attenuated HAV is useful for inducing protective immunity against HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by several nucleotide changes distributed throughout the genome, is attenuated for chimpanzees, elicits serum neutralising antibodies, and is suitable for use as an HAV vaccine. It is noted that not all the changes are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)		
Sequence	2227 AA:	
Query Match	100.0%; Score 101; DB 11; Length 2227;	
Best Local Similarity	100.0%; Pred. No. 1.4e-07;	
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 FSOAKISLFYTEHEIMKFS 20	
Db	834 FSOAKISLFYTEHEIMKFS 853	

AAW34074	standard; Protein; 2227 AA.
AAW34074	
27-APR-1998	(first entry)
Hepatitis A virus HM-175 protein sequence.	
HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection; vaccine.	
Hepatitis A virus HM-175.	
Key	Location/Qualifiers
Protein	1..23
Protein	/label= VP4
Protein	24..245
Protein	/label= VP2
Protein	246..491
Protein	/label= VP3
Protein	492..791
Protein	/label= VP1
Protein	792..980
Protein	/label= 2A
Protein	981..1087
Protein	/label= 2B
Protein	1088..1422
Protein	/label= 2C
Protein	1423..1496
Protein	/label= 3A
Protein	1497..1519
Protein	/label= 3B
Protein	1520..1738
Protein	/label= 3C
Protein	1739..2227
Protein	/label= 3D
MO9740166-A2.	
30-OCT-1997.	
18-APR-1997;	97MO-US06506.
19-APR-1996;	96US-0015642.
(USSH) US SEC DEPT HEALTH.	
(USSH) US DEPT HEALTH & HUMAN SERVICES.	
Emerson SU, Purcell RH, Raychaudhuri G;	
WPI; 1997-535850/49.	
N-PSDB; AAT93023.	
Human attenuated HAV genome containing simian HAV 2C gene - useful as vaccines against HAV infection	
Disclosure; Fig 13A-D; 66pp; English.	
This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (I) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (I); (2) a cell transfected with (I) or the RNA transcript of (I); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (I) or its RNA transcript can be used as a vaccine for preventing HAV in	

CC a mammal. (1) or the RNA transcript can also be used to stimulate
 CC the production of protective antibodies in the mammal.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 18; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||||

DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 8

AAB18607
 ID AAB18607 standard; Protein; 2227 AA.

XX AAB18607;

AC 15-JAN-2001 (first entry)

DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

OS US6113912-A.

PN 05-SEP-2000.

XX 07-JUN-1995; 95US-0475886.

PF 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

DR N-PSDB; AAA75476.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type

XX Disclosure; Fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus

CC (HAV) strain HM-174. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||||

DB 834 FSOAKISLFYTEHEIMKFS 853

AC AAB18608;
 XX 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX P-35 virus.

OS Hepatitis A virus.

PN US6113912-A.

XX 05-SEP-2000.

XX 07-JUN-1995; 95US-0475886.

PF 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

DR N-PSDB; AAA75477.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type

XX Disclosure; Columns 67-78; 72pp; English.

XX The present sequence is derived from passage 35 of a wild type

CC hepatitis A virus (HAV) strain HM-174. The resulting virus is

CC designated P-35 virus. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||||

DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 10

AAB18609
 ID AAB18609 standard; Protein; 2227 AA.

XX AAB18609;

AC 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX HAV 4380.

XX Hepatitis A virus.

OS US6113912-A.

PN 05-SEP-2000.

XX 07-JUN-1995; 95US-0475886.

```
XX 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
DR WPI: 2000-586464/55.
DR N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
PS Disclosure; Columns 93-104; 72pp; English.
XX
CC The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.
SQ Sequence 2227 AA;
XX
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOAKISLFYTEHEHMKFS 853
XX
RESULT 11
ABG31727
ID ABG31727 standard; Protein; 2227 AA.
XX
AC ABG31727;
XX
DE 29-NOV-2002 (first entry)
XX
DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX
KW Hepatitis A virus strain HM-175.
XX
PN US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-0653499.
XX
PR 07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX
DR WPI: 2002-680946/73.
DR N-PSDB; ABB52787.
XX
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection
XX
PS Disclosure; Fig 6; 71pp; English.
```

```
XX
CC The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC strain HM-175 polypeptide.
XX
SQ Sequence 2227 AA;
XX
Query Match 100.0%; Score 101; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOAKISLFYTEHEHMKFS 853
XX
RESULT 12
ABG31728
ID ABG31728 standard; Protein; 2227 AA.
XX
AC ABG31728;
XX
DE 29-NOV-2002 (first entry)
XX
DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
KW virucide; mutant; PHAV/7; mutin.
XX
OS Hepatitis A virus strain HM-175.
OS synthetic.
XX
FH Key Location/Qualifiers
FH MISC-difference 963
FT /label= Wild-type Lys substituted by Arg
FT MISC-difference 764
FT /note= "Wild-type Glu substituted by Val"
FT MISC-difference 821
FT /note= "Wild-type Asn substituted by Ser"
FT MISC-difference 1052
FT /note= "Wild-type Ala substituted by Val"
FT MISC-difference 1062
FT /note= "Wild-type Gly substituted by Ala"
FT MISC-difference 1118
FT /note= "Wild-type Lys substituted by Met"
FT MISC-difference 1151
FT /note= "Wild-type Glu substituted by Lys"
FT MISC-difference 1163
FT /note= "Wild-type Phe substituted by Ser"
FT MISC-difference 1277
FT /note= "Wild-type Val substituted by Ile"
FT MISC-difference 1500
FT /note= "Wild-type His substituted by Tyr"
FT MISC-difference 1805
FT /note= "Wild-type Asp substituted by Asn"
FT MISC-difference 1930
FT /note= "Wild-type Ser substituted by Thr"
XX
PN US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-0653499.
XX
PR 07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
```

PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
 XX WPI: 2002-680946/73.
 DR N-PSDB; ABS52788.
 XX
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection
 XX
 PS Example 3; Column 67-78; 71pp; English.
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC mutant strain HM-175/7 (pHAV/7) polypeptide.
 XX
 SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||
 DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 13

ABG31729
 ID ABG31729 standard; Protein; 2227 AA.

AC ABG31729;
 XX

DT 29-NOV-2002 (first entry)
 XX

DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
 XX

KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 KW HAV 4380.
 XX

OS Hepatitis A virus strain HM-175.
 XX

PN US6423318-B1.
 XX

PD 23-JUL-2002.
 XX

PF 31-AUG-2000; 2000US-0653499.
 XX

PR 07-JUN-1995; 95US-0475886.
 PR 17-SEP-1993; 93US-0397232.
 PR 17-SEP-1993; 93WO-US08610.
 XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX

PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
 XX

DR WPI: 2002-680946/73.
 DR N-PSDB; ABS52789.
 XX

PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection
 XX

PS Disclosure; Column 93-104; 71pp; English.
 XX

CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents an attenuated
 CC hepatitis A virus 4830 polypeptide.
 XX

SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||
 DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 14

AAE19899
 ID AAE19899 standard; Protein; 2227 AA.

AC AAE19899;
 XX

DT 18-JUN-2002 (first entry)
 XX

DE Hepatitis A virus (HAV) protein.
 XX

KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
 XX

OS Hepatitis A virus.
 XX

PN WO200213855-A2.
 XX

PD 21-FEB-2002.
 XX

PF 15-AUG-2001; 2001WO-IB01808.
 XX

PR 17-AUG-2000; 2000US-225767P.
 PR 29-AUG-2000; 2000US-229175P.
 PR 03-NOV-2000; 2000US-0705547.
 XX

PA (TRIP-) TRIPEP AB.
 XX

PI Salberg M, Hultgren C;
 XX

DR WPI: 2002-241837/29.
 DR N-PSDB; AAD31766.
 XX

XX
 XX

PT Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus
 XX

PS Claim 11; Page 82-87; 120pp; English.
 XX

CC The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein.
 XX

SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
 ||||||||||||||||
 DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 15

ABU08639
 ID ABU08639 standard; Protein; 2227 AA.
 XX

AC ABU08639;
 XX

XX

DT 03-JUN-2003 (first entry)
XX
XX Wild type human hepatitis A virus strain HM-175.
DE
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection.
XX
XX Hepatitis A virus strain HM-175.
OS
XX
XX US2002176869-A1.
PN
XX
XX 28-NOV-2002.
PD
XX
XX 29-APR-2002; 2002US-0135988.
PF
XX
XX 07-JUN-1995; 95US-0475886.
PR 31-AUG-2000; 2000US-0653499.
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 17-APR-1995; 95US-0397232.
PP
XX
XX (FUNK/) FUNKHOUSER A W.
XX (EMER/) EMERSON S U.
XX (PURC/) PURCELL R H.
XX (DHON/) D'HONDT E.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
PI
XX
XX WPI: 2003-352605/33.
DR N-PSDB: ABX93473.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PR and disease
XX
XX
XX Disclosure: Fig 6; 70pp: English.
PS
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175.
XX
XX
XX Sequence 2227 AA;
SQ

Query Match 100.0%; Score 101; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSOAKISLFYEHEHIMKES 20
|||||
834 FSOAKISLFYEHEHIMKES 853

Search completed: October 1, 2003, 09:56:42
Job time : 44.5294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds

(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101

Sequence: 1 FSOAKISLFTYEHHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_76:*

1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	1358	2 A03905	genome polypotein
2	101	100.0	2227	1 GNNYHM	genome polypotein
3	101	100.0	2227	1 GNNYMR	genome polypotein
4	101	100.0	2227	1 GNNYMK	genome polypotein
5	101	100.0	2227	1 GNNYHB	genome polypotein
6	97	96.0	852	1 GNNYHA	genome polypotein
7	88	87.1	2230	1 GNNYSA	genome polypotein
8	47	46.5	378	2 F64300	formate dehydrogen
9	43	42.6	325	2 H71869	hypothetical prote
10	43	42.6	410	2 T22413	hypothetical prote
11	42	41.6	125	2 AB1747	glycine cleavage s
12	42	41.6	125	2 A11377	glycine cleavage s
13	42	41.6	187	2 T18844	hypothetical prote
14	42	41.6	545	2 S10842	gene ND5 intron 2
15	42	41.6	1265	2 F88866	protein F5262.2 [1
16	41.5	41.1	288	2 S64243	hypothetical prote
17	41	40.6	144	2 E75128	glycine cleavage s
18	41	40.6	190	2 C90529	nuclease, lipoprot
19	41	40.6	209	2 S65300	hypothetical prote
20	41	40.6	304	2 C72324	homoserine O-succ
21	41	40.6	465	2 B97235	deacetylase/dipep
22	41	40.6	483	2 G84113	hypothetical prote
23	41	40.6	557	2 S58164	glucose-6-phosphat
24	41	40.6	820	2 T51510	hypothetical prote
25	41	40.6	1040	1 A38306	alpha-mannosidase
26	41	40.6	2431	1 MNWVSF	nonstructural poly
27	40.5	40.1	251	2 T13104	minor tail protein
28	40.5	40.1	419	2 T25577	hypothetical prote
29	40	39.6	142	2 C70059	hypothetical prote

30	40	39.6	203	2 S63633	hypothetical prote
31	40	39.6	256	2 T45895	hypothetical prote
32	40	39.6	298	2 A89009	protein T276.1 [1
33	40	39.6	333	2 T33866	hypothetical prote
34	40	39.6	450	2 C82204	heat shock protein
35	40	39.6	523	2 D83631	probable sulfate t
36	39.5	39.1	109	2 S74054	hypothetical prote
37	39.5	39.1	260	2 B71353	probable D,D-carbo
38	39.5	39.1	410	2 T23264	hypothetical prote
39	39.5	39.1	507	2 B89792	hypothetical prote
40	39	38.6	129	2 F72075	hypothetical prote
41	39	38.6	129	2 G86547	hypothetical prote
42	39	38.6	145	2 F83247	hypothetical prote
43	39	38.6	152	1 C69202	conserved hypothet
44	39	38.6	167	2 F83845	hypothetical prote
45	39	38.6	171	2 T18167	hypothetical prote

ALIGNMENTS

RESULT 1

A03905

genome polypotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Matzel Jr., J.V.; Purcell, R.H.; Felna

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A>Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polypotein

C:Keywords: coat protein; core protein; polypotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match

Best Local Similarity 100.0%; Score 101; DB 2; Length 1358;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFTYEHHEIMKFS 20

DB 834 FSOAKISLFTYEHHEIMKFS 853

RESULT 2

GNNYHM

genome polypotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core

B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A>Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-References: EMBL:M14707; NID:g329582; PIDN:AAA5465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polypotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>

C:Accession: A03904
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793
A:Accession: A03904
A:Molecule type: genomic RNA
A:Residues: 1-852 <LIN>
A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; polypeptide
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 96.0%; Score 97; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 4.7e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSOAKISLFYTEHEIMKRF 19
|||||
DB 834 FSOAKISLFYTEHEIMKRF 852

RESULT 7

GENYSA

genome polypeptide - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470

A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
A:Reference number: J01080; MUID:9111420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885

A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A:Reference number: S03965; MUID:8932168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; polypeptide
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: core protein 3A #status predicted <C3A>
F:1499-1521/Product: core protein 3B #status predicted <C3B>
F:1522-1741/Product: core protein 3C #status predicted <C3C>
F:1742-2230/Product: protein 3D #status predicted <C3D>

Query Match 87.1%; Score 88; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 4.6e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 FSOAKISLFYTEHEIMKRF 20
|||||
DB 838 FSOAKISLFYTEHEIMKRF 857

RESULT 8

F64300

formate dehydrogenase (EC 1.2.1.2) alpha chain - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: F64300

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
; Welch, C.L.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodex,
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: F64300
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-378 <BUL>
A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AA897987.1; PID:g1498763
C:Genetics:
A:Map position: REV8474-7338
C:Keywords: oxidoreductase

Query Match 46.5%; Score 47; DB 2; Length 378;
Best Local Similarity 47.1%; Pred. No. 4.7;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

4 ARISLFYTEHEIMKRF 20
|||||
DB 174 SKVTIPNTEKEIKLKN 190

RESULT 9

H71869

hypothetical protein jhp0940 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 04-Mar-2000
C:Accession: H71869

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71869
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <ARN>
A:Cross-references: GB:AE001523; GB:AE001439; NID:g4155513; PIDN:AA06516.1; PID:g415
A:Experimental source: strain J99
C:Superfamily: Helicobacter pylori hypothetical protein jhp0940

Query Match 42.6%; Score 43; DB 2; Length 325;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

5 KISLFYTEHEIMKRF 19
|||||
DB 24 KISLFYNMELVWVWF 38

RESULT 10

THIS PAGE BLANK (USPTO)

OM protein - protein search, using sw model

Title: US-09-171-432A-43

Sequence: 1 FSOAKISLFYTEHEIMKFS 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database : SwlSProt_41:4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	101	100.0	2227	1	POLG_HPAVH	P08617 hepatitis a
2	101	100.0	2227	1	POLG_HPAVL	P06441 hepatitis a
3	101	100.0	2227	1	POLG_HPVM	P13901 hepatitis a
4	97	96.0	852	1	POLG_HPAVC	P06442 hepatitis a
5	93	92.1	2226	1	POLG_HPAV2	P26560 hepatitis a
6	93	92.1	2226	1	POLG_HPAV4	P26561 hepatitis a
7	93	92.1	2226	1	POLG_HPAV8	P26582 hepatitis a
8	88	87.1	2230	1	POLG_HPAP5	P14553 simian hepea
9	47	46.5	378	1	PDHA_METJ4	P03134 methanococci
10	44	43.6	509	1	MATK_NWMD	P02055 caenorhabdit
11	43	42.6	410	1	PSD6_CAEEL	P08628 latyrus tti
12	43	42.6	506	1	MATK_LATTI	P08452 kladsjella
13	43	42.6	722	1	YC06_KLEPN	P02813 listeria in
14	42	41.6	125	1	GC5H_LISTIN	P08y412 listeria mc
15	42	41.1	125	1	GC5H_LISTMO	P53081 saccharomyc
16	41.5	41.1	288	1	NIF3_YEAST	P09w03 thermocoga
17	41	40.6	138	1	GC5H_PYRAB	P09w01 pyrococcus
18	41	40.6	304	1	META_THEMA	P59008 acinetobact
19	41	40.6	557	1	G6P1_ACICA	P21139 mus musculu
20	41	40.6	1039	1	M2C1_MOUSE	P21139 rattus norv
21	41	40.6	1040	1	M2C1_RAT	P08u411 semiliki fo
22	41	40.6	2431	1	POLN_SFV	P08u41 pyrococcus
23	40	39.6	134	1	GC5H_PYRFU	P28004 bos taurus
24	40	39.6	418	1	B2AR_BOVIN	P09f60 homo sapien
25	40	39.6	5147	1	PCIO_HUMAN	P08z30 pyrobaculum
26	39.5	39.1	347	1	HTPX_PYRAE	P10952 saccharomyc
27	39	38.6	306	1	MK16_YEAST	P07285 saccharomyc
28	39	38.6	380	1	TRPD_YEAST	P35504 homo sapien
29	39	38.6	534	1	UD15_HUMAN	P31916 euglena grac
30	39	38.6	634	1	CXK3_EUGGR	P10089 escherichia
31	39	38.6	707	1	HLV2_ECOLI	P08716 escherichia
32	39	38.6	707	1	HLV8_ECOLI	P08716 escherichia
33	39	38.6	707	1	HLV8_PROVU	P11599 proteus vni

[illegible]

ALIGNMENTS

	RESULT_1		
POLG_HPAVH	STANDARD:	PRT:	2227 AA.
AC	P06617; P06643; O81082;		
DT	01-AUG-1988 (Rel. 08, Created)		
DT	01-AUG-1988 (Rel. 08, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Genome polypeptid [Contains: Coat proteins VP1 TO VP4; Core proteins		
DE	P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase		
DE	P3D (EC 2.7.7.48)].		
OS	Hepatitis A virus (strain HM-175).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
OC	Hepatovirus.		
OX	NCB1_TaxId=12098;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-Mild type;		
RX	MEDLINE=87061253; PubMed=3023706;		
RA	Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-white A.,		
RA	Baroudy B.M.;		
RT	"Complete nucleotide sequence of wild-type hepatitis A virus:		
RT	comparison with different strains of hepatitis A virus and other		
RT	plicornaviruses.";		
RL	J. Virol. 61:50-59(1987).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-Attenuated;		
RX	MEDLINE=87175701; PubMed=3031686;		
RA	Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,		
RA	Purcell R.H.;		
RT	"Complete nucleotide sequence of an attenuated hepatitis A virus:		
RT	comparison with wild-type virus.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).		
RN	[3]		
RP	SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.		
RX	MEDLINE=85166289; PubMed=2984684;		
RA	Baroudy B.M., Ticehurst J.R., Miele T.A., Malzel J.V. Jr.,		
RA	Purcell R.H., Feinstein S.M.;		
RT	"Sequence analysis of hepatitis A virus cDNA coding for capsid		
RT	proteins and RNA polymerase."		
RL	Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).		
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +		
CC	(RNA)(N).		
CC	-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,		
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,		
CC	VP3, AND VP4.		
CC	-1- PMM: SPECIFIC ENZYMAITC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.		
CC	-1- MISCELLANEOUS: STRAIN HM-175/77 MK-5 IS ATTENUATED STAIN DERIVED		
CC	FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.		
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.		
CC	-1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT		
CC	SHOWN.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M14114; AAA45475.1; -
DR EMBL: M14707; AAA45465.1; -
DR EMBL: M14707; AAA45466.1; ALT-INT.
DR EMBL: M16632; AAA45471.1; -
DR PIR: A25981; GNNYHM.
DR PIR: A94149; GNNYMK.
DR PDB: 1HAV; 23-DEC-96.
DR MEROPS; C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSV1r.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.

FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
FT CHAIN 77 77
FT VARIANT 764 764
FT VARIANT 821 821
FT VARIANT 1052 1052
FT VARIANT 1062 1062
FT VARIANT 1118 1118
FT VARIANT 1151 1151
FT VARIANT 1163 1163
FT VARIANT 1277 1277
FT VARIANT 1500 1500
FT VARIANT 1805 1805
FT VARIANT 1930 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E235E7AEB740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTHEHEIMKFS 20
Db 834 FSOAKISLFYTHEHEIMKFS 853

RESULT 2
ID POLG_HPAVL STANDARD; PRT: 2227 AA.
AC P06441;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain LA).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12099;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=85190549; PubMed=2986127;
RA Natarajan R., Caput D., Gee W.W., Potter S.J., Renard A.,
RA Merryweather J., van Nest G., Dina D.;
RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: K02990; AAA45472.1; -
DR PIR: A03903; GNNYHR.
DR MEROPS; C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSV1r.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354BCD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTHEHEIMKFS 20
Db 834 FSOAKISLFYTHEHEIMKFS 853

RESULT 3
ID POLG_HPAVL STANDARD; PRT: 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MBB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12100;
RN [1]
RP SEQUENCE FROM N.A.


```

RX MEDLINE-86045071; PubMed-2823500;
RA Paul A.V., Tada H., der Heim R., Wissel T., Klein R., Wimmer E.,
RA Dehnardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
RL virus (isolate MB8)."
CC
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: M20773; AAA45474.1; -.
CC
CC DR MEROPS: C03.005; -.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVir.
CC DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam: PF00910; RNA_helicase; 1.
CC
CC KW Polypeptide; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC FT CHAIN 837 980 CORE PROTEIN P2A.
CC FT CHAIN 981 1087 CORE PROTEIN P2B.
CC FT CHAIN 1088 1422 CORE PROTEIN P2C.
CC FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
CC FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
CC FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
CC FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
CC FT CHAIN 2227 AA; EC983ED2A7C86349 CRC64;
CC
CC SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

```

```

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

```

```

RESULT 4
POLG_HPAVC STANDARD: PRT; 852 AA.
ID POLG_HPAVC 083741; 083742;
AC 01-JAN-1988 (Rel. 06; Created)
DT 01-JAN-1988 (Rel. 06; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,

```

```

RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA."
RL J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: M10033; AAA45470.1; -.
CC
CC DR PIR: A03904; GNNYHA.
CC KW Polypeptide; Coat protein; Core protein.
CC
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC FT CHAIN 837 >852 CORE PROTEIN P2A.
CC FT NON_TER 852 852
CC
CC SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

```

```

Query Match 96.0%; Score 97; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FSOAKISLFYTEHEIMKFS 19
DB 834 FSOAKISLFYTEHEIMKFS 852

```

```

RESULT 5
POLG_HPAV2 STANDARD: PRT; 2226 AA.
ID POLG_HPAV2
AC P26580;
DT 01-AUG-1992 (Rel. 23; Created)
DT 01-AUG-1992 (Rel. 23; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Felstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M59810; AAA45468.1; -
DR MEROPS: C03.005; -
DR InterPro: IPR004004; Calicifol. pol. hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVLr.
DR Pfam: PF00680; RNA_dep.RNA_pol. 1.
DR Pfam: PF00910; RNA_helicase. 1.
DR PRINTS: PR00918; CALICIVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. No. 2.6e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOANISLFYTEHEHMKFS 853

RESULT 6
POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID POLG_HPAV4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatitis A virus.
NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; Pubmed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cromens T., Jansen R.W.,
RT Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
(RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC or send an email to license@isb-sib.ch.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M59809; AAA45469.1; -
DR MEROPS: C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVLr.
DR Pfam: PF00680; RNA_dep.RNA_pol. 1.
DR Pfam: PF00910; RNA_helicase. 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. No. 2.6e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOANISLFYTEHEHMKFS 853

RESULT 7
POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID POLG_HPAV8
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatitis A virus.
NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; Pubmed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
RA Cromens T., Jansen R.W.,
RT Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
(RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```
CC -----
DR EMBL: M59808; AAA45467.1;
DR PDB: 1OAT; 15-MAY-00.
DR MEROPS: C03.005;
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVL.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
KM Polyprotein; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
ET CHAIN 1
ET CHAIN 23
ET CHAIN 24
ET CHAIN 245
ET CHAIN 246
ET CHAIN 491
ET CHAIN 492
ET CHAIN 794
ET CHAIN 795
ET CHAIN 900
ET CHAIN 901
ET CHAIN 1087
ET CHAIN 1088
ET CHAIN 1422
ET CHAIN 1423
ET CHAIN 1495
ET CHAIN 1496
ET CHAIN 1518
ET CHAIN 1519
ET CHAIN 1737
ET CHAIN 1738
ET CHAIN 2226
ET CHAIN 251292
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match
Best Local Similarity 92.1%; Score 93; DB 1; Length 2226;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOAKISLFYTEHEHMKFS 853

RESULT 8
POLG_HPAVS
ID FOLG_HPAVS STANDARD: PRT; 2230 AA.
AC P1453;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
XX NCBI_Taxid=12102;
XX (1)
XX SEQUENCE FROM N.A.
RX MEDLINE-89232168; PubMed-1649901;
RX Tsarev S.A., Emerson S.O., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains."
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RX Balayan M.S., Kusov Y.Y., Andjeparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chiznikov V.E., Blinov V.M., Vasilenko S.K.;
RA "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses."
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D00924; BAA00766.1;
DR EMBL: X15461; CAA33490.1;
DR PIR: A30470; GNNYSA.
DR MEROPS: C03.005;
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVL.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICVIRUSNS.
KM Polyprotein; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.
ET CHAIN 1
ET CHAIN 27
ET CHAIN 28
ET CHAIN 249
ET CHAIN 250
ET CHAIN 495
ET CHAIN 496
ET CHAIN 795
ET CHAIN 796
ET CHAIN 984
ET CHAIN 985
ET CHAIN 1091
ET CHAIN 1092
ET CHAIN 1426
ET CHAIN 1427
ET CHAIN 1498
ET CHAIN 1499
ET CHAIN 1521
ET CHAIN 1522
ET CHAIN 1741
ET CHAIN 1742
ET CHAIN 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3320E324E1F19 CRC64;

Query Match
Best Local Similarity 87.1%; Score 88; DB 1; Length 2230;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 838 FSOAKISLFYTEHEHMKFS 857

RESULT 9
FDHA_METJA
ID FDHA_METJA STANDARD: PRT; 378 AA.
AC Q60314;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase alpha chain (EC 1.2.1.2).
GN MJ0006.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
XX NCBI_Taxid=2190;
XX (1)
XX SEQUENCE FROM N.A.
RX STRAIN-JAI-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed-8688087;
RX Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D.,
RA Overbeck R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weisman J.E., Funtmann J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanne M.C.,
RA Cotton M.P., Roberts C.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN), ZINC, FAD, MAY BIND A 4FE-4S
```

```

CC CLUSTER (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U67459; AAB97987.1;
CC PIR: F64300; F64300.
CC HSSP: P07658; IAA6.
CC TIGR: MJO006;
CC DR InterPro: IPR006653; Molybdop_Fe4S4.
CC DR InterPro: IPR006656; Molybdoplerin.
CC DR InterPro: IPR006655; Prok_Mboxred.
CC Pfam: PF04879; Molybdop_Fe4S4; 1.
CC Pfam: PF00384; Molybdoplerin; 2.
CC PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; 1.
CC PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
CC DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
CC DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
CC Hypothetical protein; Oxidoreductase; Zinc; Flavoprotein; Molybdenum;
CC FAD; NAD; Iron-sulfur; 4Fe-4S; Complete proteome.
CC FT METAL 8 8 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 11 11 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 15 15 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 43 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC SO SEQUENCE 378 AA; 42050 MW; EC13060FF2609AA CRC64;
-----
Query Match 46.5%; Score 47; DB 1; Length 378;
Best Local Similarity 47.1%; Pred. No. 2.1;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
OY 4 AKISLFTTEHEIMKEFS 20
DB 174 SKVITFTEHEIKLN 190
-----
RESULT 10
MATE_NYMOD
ID MATE_NYMOD STANDARD; PRT; 509 AA.
AC 098637;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
PT Maturase K (intron maturase).
MATE.
OS Nymphaea odorata (White water lily).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Nymphaeaceae; Nymphaea.
OX NCBI_TaxID=4419;
[1]
RN SEQUENCE FROM N.A.
RP Les D.H., Schneider E.L., Padgett D.J., Soltis P.S., Soltis D.E.,
RA Zanis M.;
RA "Phylogeny, classification and floral evolution of water lilies
RT (Nymphaeaceae; Nymphaeales): a synthesis of non-molecular, rbcL, matK
RT and 18S rDNA data.";
RL Syst. Bot. 24:28-46(1999).
CC -1- FUNCTION: Probably assists in splicing chloroplast group II
CC introns (BY similarity).
CC -1- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATE
CC SUBFAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AF092988; AAD05556.1;
CC DR InterPro: IPR000442; Intron_maturase2.
CC DR InterPro: IPR002866; MatK_N.
CC Pfam: PF01348; Intron_maturase2; 1.
CC DR Pfam: PF01824; MatK_N; 1.
CC DR mRNA processing; Chloroplast.
CC SO SEQUENCE 509 AA; 60343 MW; C78FCA7301BDF373 CRC64;
-----
Query Match 43.6%; Score 44; DB 1; Length 509;
Best Local Similarity 72.7%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 9 FYTEHEIMKE 19
DB 472 FLTEHEIVSP 482
-----
RESULT 11
PSD6_CAEEL
ID PSD6_CAEEL STANDARD; PRT; 410 AA.
AC Q20585;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable 26S proteasome non-ATPase regulatory subunit 6.
GN RPN-7 OR P49C12.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RA Gardner A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
CC is involved in the ATP-dependent degradation of ubiquitinated
CC proteins (BY similarity).
CC -1- DOMAIN: CONTAINS 1 PCI DOMAIN.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: Z68227; CAA92512.1;
CC DR PIR: T22413; T22413.
CC DR Wormpep: P49C12.8; CE03368.
CC DR InterPro: IPR000717; PCI.
CC DR Pfam: PF01399; PCI; 1.
CC DR SMART: SM00088; PINT; 1.
CC KM Proteasome.
CC FT DOMAIN 205 379 PCI.
CC SO SEQUENCE 410 AA; 47583 MW; F37390A3250109EE CRC64;
-----
Query Match 42.6%; Score 43; DB 1; Length 410;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
OY 1 FSOAKISLFTTEHEIMKE 19
DB 156 FAMIRVGLFLDHLINKF 174
-----
RESULT 12
MATE_LATTT

```

```

ID  MATK_LATTI      STANDARD:      PRT:      506 AA.
AC  Q8MCR8;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DE  28-FEB-2003 (Rel. 41, Last annotation update)
GN  Maturase K (Intron maturase).
OS  Lathyrus lingulatus (Tangier pea).
OC  Chloroplast.
OC  Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC  Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OC  eurosids I: Fabales; Fabaceae; Papilionoideae; Viciae; Lathyrus.
OX  NCBI_TaxID=3862;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Steele K.P., Wojciechowski M.F.:
RT  Phylogenetic analyses of tribes Trifolieae and Viciae based on
RT  sequences of the plastid gene matK (Papilionoideae: Leguminosae).";
RT  Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: Probably assists in splicing chloroplast group II
CC  introns (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC  SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF522087; AAM82079.1; -
DR  InterPro: IPR000442; Intron_maturase2.
DR  InterPro: IPR002866; MatK_N.
DR  Pfam: PF01348; Intron_maturase2; 1.
DR  Pfam: PF01824; MatK_N; 1.
DR  MRNA processing: Chloroplast.
SQ  SEQUENCE 506 AA; 60897 MW; 6187B6AE73C9D564 CRC64;

Query Match      42.6%; Score 43; DB 1; Length 506;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY  2 SQAKISLFYTEHEIMK 19
DB  457 SEELQEFTEGELISF 474

RESULT 13
YC06_KLEPN      STANDARD:      PRT:      722 AA.
ID  YC06_KLEPN
AC  Q48452;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Putative tyrosine-protein kinase in cps region (EC 2.7.1.112) (ORF6).
OS  Klebsiella pneumoniae.
OS  Bacteria: Proteobacteria: Gammaproteobacteria: Enterobacteriales;
OC  Enterobacteriaceae: Klebsiella.
OX  NCBI_TaxID=573;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-Chedid;
RX  MEDLINE=95204345; PubMed=7896702;
RA  Arakawa Y., Wacharotayanankun R., Nagatsuka T., Ito H., Kato N.,
RA  Ohta M.:
RT  Genomic organization of the Klebsiella pneumoniae cps region
RT  responsible for serotype K2 capsular polysaccharide synthesis in the
RT  virulent strain Chedid.";
RL  J. Bacteriol. 177:1788-1796(1995).
CC  -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC  tyrosine phosphate.

```

```

CC  -1- PATHWAY: Exopolysaccharide biosynthesis.
CC  -1- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE ETX/WZC FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D21242; BAA04777.1; -
DR  InterPro: IPR003856; LPS_Wzz_MPA.
DR  Pfam: PF02706; wzz; 1.
KW  Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW  Exopolysaccharide synthesis; Transmembrane; Inner membrane;
KW  ATP-binding. 31 53 POTENTIAL.
FT  TRANSMEM 427 449
SQ  SEQUENCE 722 AA; 80400 MW; 3CAD6910AEB1C3D7 CRC64;

Query Match      42.6%; Score 43; DB 1; Length 722;
Best Local Similarity 44.4%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY  1 ESQAKISLFYTEHEIMK 18
DB  326 FRESISQLYTKHEPYRK 343

RESULT 14
GCSS_LISIN      STANDARD:      PRT:      125 AA.
ID  GCSS_LISIN
AC  Q928T3;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Probable glycine cleavage system H protein.
GN  GCVH OR LIN2519.
OS  Listeria innocua.
OC  Bacteria: Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX  NCBI_TaxID=1642;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-CLIP 11262 / Serovar 6a;
RX  MEDLINE=21537279; PubMed=11679669;
RA  Glaser P., Franjeul L., Buchrieser C., Rusnick C., Amend A.,
RA  Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA  Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA  Entian K.-D., Esli H., Garcia-del Portillo F., Garrido P.,
RA  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA  Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkhat G.,
RA  Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
RA  Nordlek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA  Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA  Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT  Comparative genomics of Listeria species.";
RL  Science 294:849-852(2001).
CC  -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC  glycine. The H protein shuttles the methylamino group of glycine
CC  from the P protein to the T protein (By similarity).
CC  -1- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC  (By similarity).
CC  -1- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC  P, T, L, and H (By similarity).
CC  -1- SIMILARITY: Belongs to the gcvH family.
CC  -1- SIMILARITY: Contains 1 lipoyl-binding domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL591983; CAD00503.1; -
DR PIR: A11377; A11377.
DR Listlist: LMO02425; -
DR HAMAP: MF_00272; -; 1.
DR InterPro: IPR002930; GCV_H.
DR Pfam: PF01597; GCV_H. 1.
DR TIGRFAMs: TIGR00527; gcvH. 1.
DR PROSITE: PS00189; LIPOYL. 1.
DR LIPOYL: Complete proteome.
KW BINDING 63
FT SEQUENCE 125 AA; 13732 MW; 94FF0032FBCFA152 CRC64;
SQ
Query Match 41.6%; Score 42; DB 1; Length 125;
Best Local Similarity 72.7%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 8 LFYTEHEHMK 18
1 | | | | | : 1
7 LFYTEHEHMK 17
By
RESULT 15
GCSH_LISMO STANDARD: PRT; 125 AA.
ID 08Y4L2:
AC 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glycine cleavage system H protein.
GN GCVH OR LMO2425.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-EGD-e / Serovar 1/2a;
MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusnlok C., Amend A.,
Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
Charbit A., Cherouani F., Couve E., de Daruvar A., Deboux P.,
Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
Entian K.-D., Fajhl H., Garcia-del Portillo F., Garrido P.,
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,
Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
Nordstedt G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
Remmel B., Rose M., Schlutener T., Simoes N., Tierrez A.,
Vazquez-Boland J.-A., Voss H., Wehland J., Cosart P.;
Comparative genomics of Listeria species.;
Science 294:849-852(2001).
-1- FUNCTION: The glycine cleavage system catalyzes the degradation of
glycine. The H protein shuttles the methylamine group of glycine
from the P protein to the T protein (By similarity).
-1- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
(By similarity).
-1- SUBUNIT: The glycine cleavage system is composed of four proteins:
P, T, L, and H (By similarity).
-1- SIMILARITY: Belongs to the gcvH family.
-1- SIMILARITY: Contains 1 lipoyl-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

CC -----
DR EMBL: AL591983; CAD00503.1; -
DR PIR: A11377; A11377.
DR Listlist: LMO02425; -
DR HAMAP: MF_00272; -; 1.
DR InterPro: IPR002930; GCV_H.
DR Pfam: PF01597; GCV_H. 1.
DR TIGRFAMs: TIGR00527; gcvH. 1.
DR PROSITE: PS00189; LIPOYL. 1.
DR LIPOYL: Complete proteome.
KW BINDING 63
FT SEQUENCE 125 AA; 13801 MW; 1B4EE1E37C8980B5 CRC64;
SQ
Query Match 41.6%; Score 42; DB 1; Length 125;
Best Local Similarity 72.7%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 8 LFYTEHEHMK 18
1 | | | | | : 1
7 LFYTEHEHMK 17
By

Search completed: October 1, 2003, 09:57:48
Job time : 8.58824 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds

(without alignments)
155,564 Million cell updates/sec

Title: us-09-171-432a-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTHEHEIMKFS 20

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_tvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	101	100.0	112	12	086534 hepatitis a
2	101	100.0	132	12	08V4L9 hepatitis a
3	101	100.0	132	12	08V4L6 hepatitis a
4	101	100.0	132	12	08V4M9 hepatitis a
5	101	100.0	132	12	08V4J9 hepatitis a
6	101	100.0	132	12	08V4N1 hepatitis a
7	101	100.0	132	12	08V4J5 hepatitis a
8	101	100.0	132	12	08V4M5 hepatitis a
9	101	100.0	132	12	08V4K9 hepatitis a
10	101	100.0	132	12	08V4N3 hepatitis a
11	101	100.0	132	12	08V4M7 hepatitis a
12	101	100.0	132	12	08V4N2 hepatitis a
13	101	100.0	132	12	08V4K7 hepatitis a
14	101	100.0	132	12	08V4M6 hepatitis a
15	101	100.0	132	12	08V4M4 hepatitis a
16	101	100.0	132	12	08V4N6 hepatitis a

17	101	100.0	132	12	08V4J7 hepatitis a
18	101	100.0	132	12	08V4L5 hepatitis a
19	101	100.0	132	12	08V4K2 hepatitis a
20	101	100.0	132	12	08V4K3 hepatitis a
21	101	100.0	132	12	08V4N7 hepatitis a
22	101	100.0	132	12	08V4K1 hepatitis a
23	101	100.0	132	12	08V4J8 hepatitis a
24	101	100.0	132	12	08V4L0 hepatitis a
25	101	100.0	132	12	08V4K4 hepatitis a
26	101	100.0	132	12	08V4M2 hepatitis a
27	101	100.0	132	12	08V4M3 hepatitis a
28	101	100.0	132	12	08V4L8 hepatitis a
29	101	100.0	132	12	08V4N5 hepatitis a
30	101	100.0	132	12	08V4N4 hepatitis a
31	101	100.0	132	12	08V4K8 hepatitis a
32	101	100.0	132	12	08V4M6 hepatitis a
33	101	100.0	132	12	08V4M8 hepatitis a
34	101	100.0	132	12	08V4L2 hepatitis a
35	101	100.0	132	12	08V4K0 hepatitis a
36	101	100.0	132	12	08V4L7 hepatitis a
37	101	100.0	132	12	08V4M1 hepatitis a
38	101	100.0	132	12	08V4M0 hepatitis a
39	101	100.0	132	12	08V4K5 hepatitis a
40	101	100.0	155	12	08B8M5 hepatitis a
41	101	100.0	155	12	08B8M4 hepatitis a
42	101	100.0	155	12	08B8M3 hepatitis a
43	101	100.0	155	12	08B8M2 hepatitis a
44	101	100.0	155	12	08B8M1 hepatitis a
45	101	100.0	155	12	08B8M0 hepatitis a

ALIGNMENTS

Result 1

ID	086534	PRELIMINARY:	PRT:	112 AA.
AC	086534:			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	VP1/2A (Fragment).			
GN	VP1/2A.			
OS	Hepatitis A virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Heptovirus.			
OX	NCBI_TaxID=12092;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=LSH/S;			
RX	MEDLINE=92348653; PubMed=1668326;			
RA	Finesch N., Cavalieri F., Garelick H., Prugnola A., Pellegrini V.,			
RA	Zuckerman A.J.;			
RT	"Characterization of a hepatitis A virus strain suitable for vaccine			
RT	production."			
RL	J. Hepatol. 13:S146-S151(1991).			
DR	EMBL: S44105; AAB22739.2; -			
FT	NON_TER			
SO	SEQUENCE			

Query Match 100.0%; Score 101; DB 12; Length 112;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTHEHEIMKFS 20
DB 75 FSQAKISLFYTHEHEIMKFS 94

RESULT 2
QY 08V4L9 PRELIMINARY: PRT: 132 AA.
AC 08V4L9;

DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-11/07/00;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386864; AAL68525.1; -.
FT NON_TER 1
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;
SQ
QY Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FSOAKISLFYTEHEIMKFS 20
92 FSOAKISLFYTEHEIMKFS 111
RESULT 3
ID 08V4L6 PRELIMINARY; PRT; 132 AA.
AC 08V4L6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-01/09/00;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386867; AAL68528.1; -.
FT NON_TER 1
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;
SQ
QY Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FSOAKISLFYTEHEIMKFS 20
92 FSOAKISLFYTEHEIMKFS 111
RESULT 4
ID 08V4M9 PRELIMINARY; PRT; 132 AA.
AC 08V4M9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-19/11/97;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386854; AAL68515.1; -.
FT NON_TER 1
FT SEQUENCE 132 AA; 15336 MW; D05235844D36C9D CRC64;
SQ
QY Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FSOAKISLFYTEHEIMKFS 20
92 FSOAKISLFYTEHEIMKFS 111
RESULT 5
ID 08V4J9 PRELIMINARY; PRT; 132 AA.
AC 08V4J9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HS-06/04/99;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL: AF386884; AAL68545.1; -.
FT NON_TER 1
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;
SQ
QY Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 FSOAKISLFYTEHEIMKFS 20
92 FSOAKISLFYTEHEIMKFS 111
RESULT 6
ID 08V4N1 PRELIMINARY; PRT; 132 AA.
AC 08V4N1;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.


```

RC STRAIN-SA-11/06/97;
RA MEDLINE-21571641; PubMed-11714971;
RX Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386852; AAL68513.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15354 MW; D05770901936C69D CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 132;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 92 FSOAKISLFYTEHEIMKFS 111
|||||

RESULT 7
OY 08VAJ5 PRELIMINARY; PRT; 132 AA.
AC 08VAJ5:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-HS-21/11/00;
RX MEDLINE-21571641; PubMed-11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386888; AAL68549.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 132;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 92 FSOAKISLFYTEHEIMKFS 111
|||||

RESULT 8
OY 08VAM5 PRELIMINARY; PRT; 132 AA.
AC 08VAM5:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-SA-11/02/98;
RX MEDLINE-21571641; PubMed-11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";

```

```

RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386858; AAL68519.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15284 MW; D8FC83D4BC38287D CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 132;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 92 FSOAKISLFYTEHEIMKFS 111
|||||

RESULT 9
OY 08VAK9 PRELIMINARY; PRT; 132 AA.
AC 08VAK9:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-HS-05/03/90;
RX MEDLINE-21571641; PubMed-11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386874; AAL68535.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 132;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 92 FSOAKISLFYTEHEIMKFS 111
|||||

RESULT 10
OY 08VAN3 PRELIMINARY; PRT; 132 AA.
AC 08VAN3:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_Taxid=12092;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-LI-18/12/95;
RX MEDLINE-21571641; PubMed-11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386850; AAL68511.1; -.
FT NON_TER 1
FT NON_TER 132
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

```

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
DB 92 FSOAKISLFYEEHEIMKFS 111

RESULT 11

ID 08V4M7 PRELIMINARY; PRT; 132 AA.
AC 08V4M7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-16/01/98;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386856; AAL68517.1; -.
FT NON_TER 1
SQ SEQUENCE 132 AA; 15284 MW; D8FC83D4BC38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
DB 92 FSOAKISLFYEEHEIMKFS 111

RESULT 12

ID 08V4N2 PRELIMINARY; PRT; 132 AA.
AC 08V4N2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SA-12/02/96;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386851; AAL68512.1; -.
FT NON_TER 1
SQ SEQUENCE 132 AA; 15268 MW; D8FC83D50C31B87D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
DB 92 FSOAKISLFYEEHEIMKFS 111

RESULT 13

ID 08V4K7 PRELIMINARY; PRT; 132 AA.
AC 08V4K7;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HS-17/01/91;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386876; AAL68537.1; -.
FT NON_TER 1
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
DB 92 FSOAKISLFYEEHEIMKFS 111

RESULT 14

ID 08V4K6 PRELIMINARY; PRT; 132 AA.
AC 08V4K6;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HS-27/02/91;
RX MEDLINE=21571641; PubMed=11714971;
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;
RT "Genetic analysis of hepatitis A virus strains recovered from the
RT environment and from patients with acute hepatitis.";
RL J. Gen. Virol. 82:2955-2963(2001).
DR EMBL; AF386877; AAL68538.1; -.
FT NON_TER 1
SQ SEQUENCE 132 AA; 15266 MW; D8FC83D4B97D3C29 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
DB 92 FSOAKISLFYEEHEIMKFS 111

RESULT 15

ID Q8V4M4 PRELIMINARY: PRT; 132 AA.
 AC Q8V4M4:
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis A virus.
 OS Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SA-22/03/99;
 RX MEDLINE=21571641; PubMed=11714971;
 RA Pina S., Buti M., Jordi R., Clemente-Casares P., Jofre J., Girones R.;
 RT "Genetic analysis of hepatitis A virus strains recovered from the
 environment and from patients with acute hepatitis.";
 J. Gen. Virol. 82:2955-2963(2001).
 EMBL; AF386859; AAL68520.1; -.
 NON_TER 1
 FT 132 132
 SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

 Query Match 100.0%; Score 101; DB 12; Length 132;
 Best Local Similarity 100.0%; Pred. No. 7.5e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 FSOAKISLPTTEHEIMKFS 20
 DB 92 FSOAKISLPTTEHEIMKFS 111

Search completed: October 1, 2003, 10:02:38
 Job time : 35.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds
(without alignments)
66,600 Million cell updates/sec

Title: US-09-171-432A-43

Perfect score: 101

Sequence: 1 FSOAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6	5516630-2
2	101	100.0	2227	3	US-08-475-886-2
3	101	100.0	2227	3	US-08-475-886-4
4	101	100.0	2227	3	US-08-475-886-6
5	101	100.0	2227	3	US-08-397-232-2
6	101	100.0	2227	3	US-08-397-232-4
7	101	100.0	2227	3	US-09-171-387-2
8	101	100.0	2227	4	US-09-653-499-2
9	101	100.0	2227	4	US-09-653-499-4
10	101	100.0	2227	4	US-09-653-499-6
11	43	40.6	171	4	US-09-134-001C-5205
12	41	40.6	165	4	US-09-107-532A-7030
13	41	40.6	173	4	US-09-585-228-4
14	41	40.6	2431	1	US-07-920-281C-2
15	41	40.6	2431	3	US-08-466-277-2
16	40	39.6	173	4	US-09-585-228-2
17	40	39.6	240	4	US-09-328-352-7537
18	40	39.6	243	4	US-09-252-991A-17307
19	39	38.6	148	4	US-09-198-452A-491
20	39	38.6	170	4	US-09-198-119C-89
21	39	38.6	289	4	US-09-305-856B-8
22	39	38.6	289	5	PCT-US92-00282-11
23	39	38.6	340	4	US-09-107-532A-3720
24	39	38.6	455	4	US-09-198-452A-372
25	39	38.6	530	4	US-09-252-991A-23666
26	39	38.6	668	4	US-09-252-991A-21541
27	39	38.6	860	1	US-08-092-817-4

28	39	38.6	860	4	US-08-485-128-4	Sequence 4, Appl1
29	39	38.6	860	4	US-09-804-778A-8	Sequence 8, Appl1
30	39	38.6	1410	2	US-08-470-058-4	Sequence 4, Appl1
31	39	38.6	1410	3	US-09-037-188-4	Sequence 4, Appl1
32	39	38.6	1410	3	US-09-285-310-4	Sequence 4, Appl1
33	38	37.6	265	1	US-07-958-551-2	Sequence 2, Appl1
34	38	37.6	265	1	US-08-129-610-7	Sequence 7, Appl1
35	38	37.6	265	1	US-08-129-609A-7	Sequence 7, Appl1
36	38	37.6	265	1	US-08-455-313-7	Sequence 7, Appl1
37	38	37.6	265	1	US-08-475-924-2	Sequence 2, Appl1
38	38	37.6	265	2	US-08-657-579A-2	Sequence 2, Appl1
39	38	37.6	265	3	US-09-224-023-7	Sequence 7, Appl1
40	38	37.6	265	5	PCT-US94-07887-7	Sequence 7, Appl1
41	38	37.6	390	4	US-09-328-352-7807	Sequence 7807, Ap
42	38	37.6	707	3	US-08-772-270A-4	Sequence 4, Appl1
43	38	37.6	707	4	US-09-062-126-4	Sequence 4, Appl1
44	37.5	37.1	86	4	US-09-328-352-6252	Sequence 6252, Ap
45	37.5	37.1	378	4	US-09-107-532A-6500	Sequence 6500, Ap

ALIGNMENTS

```
RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHRIG M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match      100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 FSOAKISLFTYEHEIMKFS 20
Db      1071 FSOAKISLFTYEHEIMKFS 1090

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SOZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262052
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 3
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1993-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

PLT 4
08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1993-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 5
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 6
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20
|||||
Db 834 FSOAKISLFTTEHEIMKFS 853

RESULT 7
US-09-171-387-2

; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:

APPLICANT: RAYCHAUDHURI, GOPA;

EMERSON, SUZANNE, U.;

PURCELL, ROBERT, H.;

TITLE OF INVENTION: SIMIAN-HUMAN HAV

HAVING A CHIMERIC 2C PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESSES:

ADDRESSER: MORGAN & FINNIGAN, L.L.P.

STREET: 345 PARK AVENUE

CITY: NEW YORK

STATE: NEW YORK

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/171,387

FILING DATE: 24-Mar-1999

APPLICATION NUMBER: PCT/US97/06506

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US60/015,642

FILING DATE: 19-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: William S. Feiler

REGISTRATION NUMBER: 26,728

REFERENCE/DOCKET NUMBER: 2026-4229US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 2

SEQUENCE CHARACTERISTICS:

LENGTH: 2227 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-171-387-2

OY 1 FSOAKISLFTTEHEIMKFS 20
|||||
Db 834 FSOAKISLFTTEHEIMKFS 853

RESULT 8
US-09-653-499-2

; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

Query Match 100.0%; Score 101; DB 3; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.3e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2

; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
; US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20
|||||
Db 834 FSOAKISLFTTEHEIMKFS 853

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

CURRENT FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: 08/475,886

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: 07/947,338

PRIOR FILING DATE: 1992-09-18

PRIOR APPLICATION NUMBER: 08/397,232

PRIOR FILING DATE: 1995-03-10

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 2227

TYPE: PRT

ORGANISM: Attenuated HAV (Pass 35), strain HM-175

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20
|||||
Db 834 FSOAKISLFTTEHEIMKFS 853

RESULT 10
US-09-653-499-6

; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

Query Match 100.0%; Score 101; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.3e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

: PRIOR APPLICATION NUMBER: 08/475,886
: PRIOR FILING DATE: 1999-08-09
: PRIOR APPLICATION NUMBER: 07/947,338
: PRIOR FILING DATE: 1992-09-18
: PRIOR APPLICATION NUMBER: 08/397,232
: PRIOR FILING DATE: 1995-03-10
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 2227
: TYPE: PRF
: ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
      |||
      834 FSOAKISLFYEEHEIMKFS 853

RESULT 11
US-09-134-001C-5205
: Sequence 5205, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: FILE REFERENCE: CTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5205
: LENGTH: 171
: TYPE: PRF
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5205

Query Match
Best Local Similarity 42.6%; Score 43; DB 4; Length 171;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SOAKISLFYEEHE 15
      |||
      93 SOAKISLFYEEHEQ 106

DB

RESULT 12
US-09-107-532A-7030
: Sequence 7030, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: PC
```

```

: OPERATING SYSTEM: <Unknown>
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532A
: FILING DATE: 30-Jun-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085,598
: FILING DATE: 14 May 1998
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Arinello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)893-5007
: TELEFAX: (781)893-8277
: INFORMATION FOR SEQ ID NO: 7030:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 165 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: YES
: ORIGINAL SOURCE:
: ORGANISM: Enterococcus faecium
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (B) LOCATION 1...165
: SEQUENCE DESCRIPTION: SEQ ID NO: 7030:
US-09-107-532A-7030

Query Match
Best Local Similarity 40.6%; Score 41; DB 4; Length 165;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KISLFYEEHEIMK 18
      |||
      1 KYSVYIREREMK 14

DB

RESULT 13
US-09-585-228-4
: Sequence 4, Application US/09585228
: Patent No. 6531576
: GENERAL INFORMATION:
: APPLICANT: Piddington, Christopher S.
: APPLICANT: West, James W.
: APPLICANT: Holly, Richard D.
: APPLICANT: Burkhead, Steven K.
: TITLE OF INVENTION: FOUR-HELICAL BUNDLE PROTEIN ZSIG81
: FILE REFERENCE: 99-13
: CURRENT APPLICATION NUMBER: US/09/585,228
: CURRENT FILING DATE: 2000-06-01
: EARLIER APPLICATION NUMBER: US 60/137,057
: EARLIER FILING DATE: 1999-06-01
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 173
: TYPE: PRF
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: (1)...(17)
US-09-585-228-4

Query Match
Best Local Similarity 40.6%; Score 41; DB 4; Length 173;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
      |||
      1 FSOAKISLFYEEHEIMKFS 20
```


Db 148 WAQSLISLFQALRHLDMSS 167

RESULT 14

US-07-920-281C-2
Sequence 2, Application US/07920281C

Patent No. 5739026

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

APPLICANT: Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

TITLE OF INVENTION: Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/920,281C

FILING DATE: 13-AUG-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2431 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-920-281C-2

Query Match 40.6%; Score 41; DB 1; Length 2431;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEH 14

Db 1072 FSAFKVSLYENNH 1085

RESULT 15

US-08-466-277-2
Sequence 2, Application US/08466277

Patent No. 6190666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

APPLICANT: Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

TITLE OF INVENTION: Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch & Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,277

FILING DATE: 06-Jun-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: 07/920,281

APPLICATION DATA:

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2431 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-466-277-2

Query Match 40.6%; Score 41; DB 3; Length 2431;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEEH 14

Db 1072 FSAFKVSLYENNH 1085

Search completed: October 1, 2003, 10:06:30

Job time: 13.7059 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using SW model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-43
Perfect score: 101
Sequence: 1 FSOAKISLFYTEHEIMKFS 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	US-10-272-459-41	Sequence 41, Appl
3	101	100.0	2227	US-09-929-955-12	Sequence 12, Appl
4	101	100.0	2227	US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	US-10-135-988-4	Sequence 4, Appl
7	101	100.0	2227	US-10-135-988-6	Sequence 6, Appl
8	47	46.5	518	US-09-253-088-6	Sequence 6, Appl
9	47	46.5	816	US-09-253-088-5	Sequence 5, Appl
10	41	40.6	1336	US-10-224-999A-3482	Sequence 3482, Ap
11	41	40.6	2431	US-09-901-106-2	Sequence 2, Appl
12	40	39.6	173	US-10-199-672-506	Sequence 506, App
13	40	39.6	173	US-10-187-749-506	Sequence 506, App
14	40	39.6	173	US-10-194-457-506	Sequence 506, App
15	40	39.6	173	US-10-184-642-506	Sequence 506, App

16	40	39.6	173	12	US-10-196-747-506	Sequence 506, App
17	40	39.6	173	12	US-10-173-689-506	Sequence 506, App
18	40	39.6	173	12	US-10-173-690-506	Sequence 506, App
19	40	39.6	173	12	US-10-173-691-506	Sequence 506, App
20	40	39.6	173	12	US-10-173-692-506	Sequence 506, App
21	40	39.6	173	12	US-10-173-693-506	Sequence 506, App
22	40	39.6	173	12	US-10-173-694-506	Sequence 506, App
23	40	39.6	173	12	US-10-173-695-506	Sequence 506, App
24	40	39.6	173	12	US-10-173-696-506	Sequence 506, App
25	40	39.6	173	12	US-10-173-697-506	Sequence 506, App
26	40	39.6	173	12	US-10-173-698-506	Sequence 506, App
27	40	39.6	173	12	US-10-173-699-506	Sequence 506, App
28	40	39.6	173	12	US-10-173-700-506	Sequence 506, App
29	40	39.6	173	12	US-10-173-701-506	Sequence 506, App
30	40	39.6	173	12	US-10-173-702-506	Sequence 506, App
31	40	39.6	173	12	US-10-173-703-506	Sequence 506, App
32	40	39.6	173	12	US-10-173-704-506	Sequence 506, App
33	40	39.6	173	12	US-10-173-705-506	Sequence 506, App
34	40	39.6	173	12	US-10-173-706-506	Sequence 506, App
35	40	39.6	173	12	US-10-173-707-506	Sequence 506, App
36	40	39.6	173	12	US-10-173-708-506	Sequence 506, App
37	40	39.6	173	12	US-10-173-709-506	Sequence 506, App
38	40	39.6	173	12	US-10-173-710-506	Sequence 506, App
39	40	39.6	173	12	US-10-173-711-506	Sequence 506, App
40	40	39.6	173	12	US-10-173-712-506	Sequence 506, App
41	40	39.6	173	12	US-10-173-713-506	Sequence 506, App
42	40	39.6	173	12	US-10-173-714-506	Sequence 506, App
43	40	39.6	173	12	US-10-173-715-506	Sequence 506, App
44	40	39.6	173	12	US-10-173-716-506	Sequence 506, App
45	40	39.6	173	12	US-10-173-717-506	Sequence 506, App

ALIGNMENTS

RESULT 1
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHONANT, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P11955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match 100.0%; Score 101; DB 15; Length 352;
Best Local Similarity 100.0%; Pred No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
Db 206 FSOAKISLFYTEHEIMKFS 225

RESULT 2
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHONANT, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```

; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFTYEEHEIMKFS 20
       |||
       834 FSOAKISLFTYEEHEIMKFS 853

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03/229,175
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
09-929-955-12

Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFTYEEHEIMKFS 20
       |||
       834 FSOAKISLFTYEEHEIMKFS 853

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFTYEEHEIMKFS 20
       |||
       834 FSOAKISLFTYEEHEIMKFS 853

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFTYEEHEIMKFS 20
       |||
       834 FSOAKISLFTYEEHEIMKFS 853

RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

```

; PRT
; ORGANISM: group B streptococcus
; US-09-252-088-6
Query Match
Best Local Similarity 100.0%; Score 101; DB 14; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 7
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: PUNKHUSER, ANN M
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 2026426203
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
Query Match
Best Local Similarity 100.0%; Score 101; DB 14; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 8
US-09-252-088-6
Sequence 6, Application US/09252088
Publication No. US20030031682A1
GENERAL INFORMATION:
APPLICANT: BRODEUR, Bernard R.
APPLICANT: RIOUX, Clement
APPLICANT: BOYER, Martine
APPLICANT: CHARLEBOIS, Isabelle
APPLICANT: HAMEL, Jose
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
FILE REFERENCE: 8331-9002
CURRENT APPLICATION NUMBER: US/09/252,088
CURRENT FILING DATE: 1999-02-18
EARLIER APPLICATION NUMBER: US/60/075,425
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 518
```

```

; PRT
; ORGANISM: group B streptococcus
; US-09-252-088-6
Query Match
Best Local Similarity 46.5%; Score 47; DB 11; Length 518;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 371 FSOAKISLFYTEHEIMKFS 390

RESULT 9
US-09-252-088-5
Sequence 5, Application US/09252088
Publication No. US20030031682A1
GENERAL INFORMATION:
APPLICANT: BRODEUR, Bernard R.
APPLICANT: RIOUX, Clement
APPLICANT: BOYER, Martine
APPLICANT: CHARLEBOIS, Isabelle
APPLICANT: HAMEL, Jose
APPLICANT: MARTIN, Denis
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
FILE REFERENCE: 8331-9002
CURRENT APPLICATION NUMBER: US/09/252,088
CURRENT FILING DATE: 1999-02-18
EARLIER APPLICATION NUMBER: US/60/075,425
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 816
TYPE: PRT
ORGANISM: group B streptococcus
US-09-252-088-5
Query Match
Best Local Similarity 45.0%; Score 47; DB 11; Length 816;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 669 FSOAKISLFYTEHEIMKFS 688

RESULT 10
US-10-224-999A-3482
Sequence 3482, Application US/10224999A
Publication No. US20030171318A1
GENERAL INFORMATION:
APPLICANT: Myriad Genetics, Inc.
APPLICANT: Morham, Scott
APPLICANT: Zavitz, Kenton
APPLICANT: Hobden, Adrian
TITLE OF INVENTION: Composition and Method for Treating Viral Infection
FILE REFERENCE: 5004.01
CURRENT APPLICATION NUMBER: US/10/224,999A
CURRENT FILING DATE: 2003-03-03
PRIOR APPLICATION NUMBER: US 60/313,695
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 3484
SOFTWARE: Patentin version 3.1
SEQ ID NO 3482
LENGTH: 1336
TYPE: PRT
ORGANISM: Semliki Forest virus
US-10-224-999A-3482
Query Match
Best Local Similarity 40.6%; Score 41; DB 12; Length 1336;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

OY 1 FSOAKISLFYTEEH 14
DB 1072 FSAKPSLVYENNH 1085

RESULT 11

US-09-901-106-2
Sequence 2, Application US/09901106
Patent No. US20020151067A1
GENERAL INFORMATION:
APPLICANT: Garoff, Henrik
Liljestrom, Peter
TITLE OF INVENTION: DNA Expression Systems Based on
Alphaviruses
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/901.106
FILING DATE: 10-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/920.281C
FILING DATE: 13-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 828-103P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-241-1300
TELEFAX: 703-241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
09-901-106-2

Query Match 40.6%; Score 41; DB 10; Length 2431;
Best Local Similarity 50.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEEH 14
DB 1072 FSAKPSLVYENNH 1085

RESULT 12

US-10-199-672-506
Sequence 506, Application US/10199672
Publication No. US20030148442A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
Chen, Jian
APPLICANT: Desnoyers, Luc
Goddard, Audrey
APPLICANT: Godowski, Paul J.
Gurney, Austin L.
APPLICANT: Pan, James

APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/199.672
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: US/10/052.586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 506
LENGTH: 173
TYPE: PRT
ORGANISM: Homo Sapien

US-10-199-672-506

Query Match 39.6%; Score 40; DB 12; Length 173;
Best Local Similarity 35.0%; Pred. No. 61;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEEHIMKFS 20
DB 148 WAOSLVSLFQALRHLMRSS 167

RESULT 13

US-10-187-749-506
Sequence 506, Application US/10187749
Publication No. US20030135036A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
Chen, Jian
APPLICANT: Desnoyers, Luc
Goddard, Audrey
APPLICANT: Godowski, Paul J.
Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/187.749
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18

```
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
10-187-749-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLMRSS 167

RESULT 14
US-10-194-457-506
; Sequence 506, Application US/10194457
; Publication No. US20030153037A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC296
; CURRENT APPLICATION NUMBER: US/10/194,457
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Prior Application data removed - See file wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-194-457-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLMRSS 167

RESULT 15
US-10-184-642-506
; Sequence 506, Application US/10184642
; Publication No. US20030157635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430RIC194
; CURRENT APPLICATION NUMBER: US/10/184,642
; CURRENT FILING DATE: 2002-06-27
; PRIOR APPLICATION removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 506
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-642-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLMRSS 167
```

Search completed: October 1, 2003, 10:37:51
Job time : 25.1176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432A-44
Perfect score: 104
Sequence: 1 KVNPRHGMLEIFRANSKD 20

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A.Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/genesep-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/genesep-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	20	AAW42927	Immunogenic Hepatitis A virus
2	104	100.0	21	AAW42927	Immunogenic Hepatitis A virus
3	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
4	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
5	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
6	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
7	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
8	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus
9	104	100.0	22	AAW42927	Immunogenic Hepatitis A virus

10	104	100.0	22	ABG31728	Hepatitis A virus
11	104	100.0	22	ABG31729	Attenuated Hepatitis A virus
12	104	100.0	22	ABE19899	Hepatitis A virus
13	104	100.0	22	ABU08639	Wild type human he
14	104	100.0	22	ABU08640	Attenuated (pass3
15	104	100.0	22	ABU08641	Immunogenic Hepati
16	53	51.0	20	AAW42928	Synthetic HAV P2A
17	47	45.2	38	ABW5661	Lactococcus lactis
18	47	45.2	38	ABW5661	Novel human secret
19	45	43.3	19	AAU32447	Arbidopais thalia
20	45	43.3	20	AAU32447	Arbidopais thalia
21	45	43.3	21	AAU32447	Arbidopais thalia
22	45	43.3	22	AAU32447	Arbidopais thalia
23	45	43.3	23	AAU32447	Arbidopais thalia
24	45	43.3	24	AAU32447	Arbidopais thalia
25	45	43.3	25	AAU32447	Arbidopais thalia
26	44	42.3	10	ABG11781	Novel human diagno
27	44	42.3	10	ABG11781	Human succinyl-Coe
28	44	42.3	10	ABG11781	Human SCSH-2 prote
29	44	42.3	10	ABG11781	Human human secret
30	44	42.3	10	ABG11781	Arbidopais thalia
31	43.5	41.8	26	AAU37721	Arbidopais thalia
32	43.5	41.8	26	AAU37721	Arbidopais thalia
33	43.5	41.8	26	AAU37721	Arbidopais thalia
34	43.5	41.8	26	AAU37721	Arbidopais thalia
35	43.5	41.8	26	AAU37721	Arbidopais thalia
36	43.5	41.8	26	AAU37721	Arbidopais thalia
37	43.5	41.8	26	AAU37721	Arbidopais thalia
38	43.5	41.8	26	AAU37721	Arbidopais thalia
39	43.5	41.8	26	AAU37721	Arbidopais thalia
40	43	41.3	98	AAU37719	Human secreted pro
41	43	41.3	98	AAU37719	Human PRO725 prote
42	43	41.3	98	AAU37719	Human angiogenesis
43	43	41.3	98	AAU37719	Human PRO725 prote
44	43	41.3	98	AAU37719	Human PRO725 prote
45	43	41.3	98	AAU37719	Human PRO725 prote

ALIGNMENTS

RESULT 1	AAW42927	standard; peptide; 20 AA.
ID	AAW42927	
AC	AAW42927	
XX		
DT	28-APR-1998	(first entry)
XX		
DE	Immunogenic Hepatitis A virus peptide YK-1327.	
XX		
KW	Immunogenic peptide; immunogenic epitope; P2A protein;	
XX	Immune response; antibody.	
OS	Synthetic.	
OS	Hepatitis A virus.	
PN	WO9740147-A1.	
XX		
PD	30-OCT-1997.	
XX		
PF	18-APR-1997;	97WO-US06891.
XX		
PR	19-APR-1996;	96US-0015644.
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX		
PI	Fields HA, Khudyakov YE;	
XX		
DR	WPI; 1997-535831/49.	
XX		
PT	Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an	
PT	Immune response to HAV in a mammal or to detect the presence of	

PT antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 922-941, and has a reactivity of 31.3% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
SQ Sequence 20 AA:
Query Match 100.0%; Score 104; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KVNPPHGMGLDEETIAANSKD 20
1 KVNPPHGMGLDEETIAANSKD 20
Db 1 KVNPPHGMGLDEETIAANSKD 20
RESULT 2
ID AAB69444 standard; Peptide; 21 AA.
XX
AC AAB69444;
XX
DT 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 44.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KM antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000MO-US19267.
XX
PS 15-JUL-1999; 99US-0144412.
XX
PI (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 96; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
SQ Sequence 21 AA:
Query Match 100.0%; Score 104; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.6e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KVNPPHGMGLDEETIAANSKD 20
1 KVNPPHGMGLDEETIAANSKD 20
Db 1 KVNPPHGMGLDEETIAANSKD 20
RESULT 3
ID AAP60066 standard; Protein; 2227 AA.
XX
AC AAP60066;
XX
DT 25-MAR-2003 (updated)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of viral L434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.
XX
KM Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
FH Key
FH Location/Qualifiers
FT 1..245
FT /label= P1.1A
FT Region
FT 246..491
FT /label= 1B
FT Region
FT 492..836
FT /label= 1C
FT Region
FT 837..980
FT /label= P2.2A
FT Region
FT 981..1076
FT /label= 2B
FT Region
FT 1077..1422
FT /label= 2C
FT Region
FT 1423..1484
FT /label= P3.3A
FT Region
FT 1485..1507
FT /label= 3B
FT Region
FT 1508..1678
FT /label= 3C
FT Region
FT 1679..2227
FT /label= 3D
XX
PN EPI99480-A.
XX
PD 29-OCT-1986.
XX
PF 03-APR-1986; 86EP-0302465.
XX
PS 03-APR-1985; 85US-0719329.
XX
PI (CHIR) CHIRON CORP.
XX
PI Dina D, Potter SJ, Vannest GA, Caput D;
XX
DR WPI; 1986-286213/44.
DR N-PSDB; AAN60080.
XX
PT Hepatitis A virus nucleotide sequence and polypeptide - and use
PT in produ. of vaccines and diagnostic probes
XX
PS Claim 5; Fig 1; 18pp; English.

XX		AAN60080 and oligonucleotide fragments are useful in detection of
CC		hepatitis A virus; transformed hosts may be used for expression of
CC		polypeptides and fragments useful in vaccines without risk of
CC		infection by the virus or in prodn. of particles which are capable
CC		of inducing immunocompetent B cells for passive immunotherapy. Pref.
CC		epitope is derived from AAs 445-657 or 792-848 of the HAV
CC		polypeptide sequence (AAP60066).
CC		(Updated on 25-MAR-2003 to correct PA field.)
XX		
SQ	Sequence	2227 AA:
Query Match	100.0%;	Score 104; DB 7; Length 2227;
Best Local Similarity	100.0%;	Pred. No. 2,8e+08;
Matches	20; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY		
DB	1 KVNPHGMLDEETIAANSKD 20	
	922 KVNPHGMLDEETIAANSKD 941	
RESULT 4		
ID	AAR05697	
XX	AAR05697 standard; protein; 2227 AA.	
AC	AAR05697;	
XX		
DT	25-MAR-2003 (updated)	
DT	15-AUG-1990 (first entry)	
XX		
DE	Attenuated hepatitis A virus.	
XX		
KW	Hepatitis A virus; vaccine; attenuated.	
XX		
OS	Hepatitis A virus, strain HM-175.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..23
FT	/label-VP4 = 1A	
FT	Region	24..245
FT	/label-VP2 = 1B	
FT	Region	246..491
FT	/label-VP3 = 1C	
FT	Region	492..791
FT	/label-VP1 = 1D	
FT	Region	792..980
FT	/label-2A	
FT	Region	981..1087
FT	/label-2B	
FT	Region	1088..1422
FT	/label-2C	
FT	Region	1423..1496
FT	/label-3A	
FT	Region	1497..1519
FT	/label-3B = VPg	
FT	Region	1520..1738
FT	/label-3C	
FT	Region	1739..2227
FT	/label-3D	
XX		
PN	US4894228-A.	
XX		
PD	16-JAN-1990.	
XX		
PE	12-JUL-1988;	88US-0217824.
XX		
PR	12-JUL-1988;	88US-0217824.
XX	12-JUL-1988;	88US-0652967.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICE.	
PI	Purcell RH, Titchhurst JR, Cohen JL, Emerson SU, Feinstein SW;	
TI	Damer RJ, Cust ID:	

XX	WP1: 1990-075557/10.
DR	N-PSDB; AAQ03512.
XX	
PT	Vaccine against hepatitis A virus infection - comprises novel
PT	attenuated hepatitis A virus strain.
XX	
PS	Claim 1; Fig 1; 18pp; English.
XX	
CC	The attenuated HAV is useful for inducing protective immunity against
CC	HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC	several nucleotide changes distributed throughout the genome, is
CC	attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC	suitable for use as an HAV vaccine. It is noted that not all the changes
CC	are necessary for attenuation and use as a vaccine.
CC	(Updated on 25-MAR-2003 to correct PA field.)
CC	(Updated on 25-MAR-2003 to correct PI field.)
XX	
SO	Sequence 2227 AA;
QY	1 KVNPHGMLDEETANSKD 20
DB	922 KVNPHGMLDEETANSKD 941
XX	
XX	AAW34074:
XX	27-APR-1998 (first entry)
XX	Hepatitis A virus HM-175 protein sequence.
XX	HA; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW	infection; vaccine.
KW	
OS	Hepatitis A virus HM-175.
XX	
XX	Key
XX	Location/Qualifiers
XX	1..23
XX	/label= VP4
XX	24..245
XX	/label= VP2
XX	246..491
XX	/label= VP3
XX	492..791
XX	/label= VP1
XX	792..980
XX	/label= 2A
XX	981..1087
XX	/label= 2B
XX	1088..1422
XX	/label= 2C
XX	1423..1496
XX	/label= 3A
XX	1497..1519
XX	/label= 3B
XX	1520..1738
XX	/label= 3C
XX	1739..2227
XX	/label= 3D
XX	
XX	MO9740166-A2.
XX	
XX	30-OCT-1997.
XX	
XX	18-APR-1997; 97WO-US06506.

XX 19-APR-1996; 960S-0015642.
PR (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Emerson SU, Purcell RH, Raychaudhuri G;
XX WPI; 1997-535850/49.
DR N-PSDB; AAT93023.
XX
PT Human attenuated HAV genome containing simian HAV 2C gene - useful
PT as vaccines against HAV infection
PS Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (1) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX
SQ Sequence 2227 AA;
OY Query Match 100.0%; Score 104; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 KVNPHGMLDLEETANSKD 20
922 KVNPHGMLDLEETANSKD 941
OY RESULT 6
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
XX
XX AAB18607;
XX
XX 15-JAN-2001 (first entry)
XX
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX
XX Hepatitis A virus.
XX
XX US613912-A.
XX
XX PD 05-SEP-2000.
XX
XX PF 07-JUN-1995; 950S-0475886.
XX
XX PR 18-SEP-1992; 920S-0947338.
XX PR 17-SEP-1993; 93MO-US08610.
XX PR 10-MAR-1995; 950S-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
XX
XX

DR N-PSDB; AAA75476.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
XX
PS Disclosure; Fig 6A-K; 72pp; English.
XX
XX The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
OY Query Match 100.0%; Score 104; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 KVNPHGMLDLEETANSKD 20
922 KVNPHGMLDLEETANSKD 941
OY RESULT 7
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
XX
XX AAB18608;
XX
XX 15-JAN-2001 (first entry)
XX
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX P-35 virus.
XX
XX Hepatitis A virus.
XX
XX US613912-A.
XX
XX PD 05-SEP-2000.
XX
XX PF 07-JUN-1995; 950S-0475886.
XX
XX PR 18-SEP-1992; 920S-0947338.
XX PR 17-SEP-1993; 93MO-US08610.
XX PR 10-MAR-1995; 950S-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
XX
XX N-PSDB; AAA75477.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
XX
PS Disclosure; Columns 67-78; 72pp; English.
XX
XX The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
XX

```

Query Match      100.0%; Score 104; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVNFPFGMDLEEIANSKD 20
      |||
DB      922 KVNFPFGMDLEEIANSKD 941

RESULT 8
AAB18609
ID      AAB18609 standard; Protein; 2227 AA.
XX
AC      AAB18609;
XX
DT      15-JAN-2001 (first entry)
XX
DE      Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX
KW      HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX      HAV 4380.
XX
OS      Hepatitis A virus.
XX
PN      US6113912-A.
XX
PD      05-SEP-2000.
XX
PF      07-JUN-1995; 95US-0475886.
XX
PR      18-SEP-1992; 92US-0947338.
XX
PR      17-SEP-1993; 93WO-US08610.
XX
PR      10-MAR-1995; 95US-0397232.
XX
PA      (USSR ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI      D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
DR      WPI; 2000-586464/55.
XX
DR      N-PSDB; AAA75478.
XX
PT      Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX      line useful as vaccine for protecting humans against hepatitis A virus
XX      infection, has modified genome compared to wild type
XX
PS      Disclosure; Columns 93-104; 72pp; English.
XX
DE      The present sequence is derived from a live attenuated hepatitis A
XX      virus (HAV) of the invention, designated HAV 4380. The sequence is
XX      produced by modifying wild type HAV strain HM-174. The HAV of the
XX      invention are adapted to growth in the human fibroblast-like cell
XX      line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
XX      appropriate attenuation. It is useful as a live vaccine for prophylaxis
XX      of hepatitis A in humans and other primates.
XX
SQ      Sequence 2227 AA;

Query Match      100.0%; Score 104; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVNFPFGMDLEEIANSKD 20
      |||
DB      922 KVNFPFGMDLEEIANSKD 941

RESULT 9
ABG31727
ID      ABG31727 standard; Protein; 2227 AA.
XX
AC      ABG31727;
XX

```

```

DT      29-NOV-2002 (first entry)
XX
DE      Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX
KW      Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX
XX      Hepatitis A virus strain HM-175.
XX
OS      US6423318-B1.
XX
PN      23-JUL-2002.
XX
PD      31-AUG-2000; 2000US-0653499.
XX
PF      07-JUN-1995; 95US-0475886.
XX
PR      17-SEP-1993; 93US-0397232.
XX
PR      17-SEP-1993; 93WO-US08610.
XX
PA      (USSR ) US DEPT HEALTH & HUMAN SERVICES.
XX      (SMIR ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI      Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX
DR      WPI; 2002-680946/73.
XX
DR      N-PSDB; ABS52787.
XX
PT      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX      in MRC-5 cells, useful for preparing a vaccine against HAV infection
XX
PS      Disclosure; Fig 6; 71pp; English.
XX
DE      The invention relates to a polynucleotide which encodes a hepatitis A
XX      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX      line). The polynucleotide is useful for preparing a vaccine against
XX      hepatitis A virus infection. This sequence represents a hepatitis A virus
XX      strain HM-175 polypeptide.
XX
SQ      Sequence 2227 AA;

Query Match      100.0%; Score 104; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KVNFPFGMDLEEIANSKD 20
      |||
DB      922 KVNFPFGMDLEEIANSKD 941

RESULT 10
ABG31728
ID      ABG31728 standard; Protein; 2227 AA.
XX
AC      ABG31728;
XX
DT      29-NOV-2002 (first entry)
XX
DE      Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
KW      Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
XX      virucide; mutant; pHAV/7; muteln.
XX
OS      Hepatitis A virus strain HM-175.
XX
OS      Synthetic.
XX
FH      Key Location/Qualifiers
FT      Misc-difference 963
FT      /label= Wild-type Lys substituted by Arg
FT      Misc-difference 764
FT      /note= "Wild-type Glu substituted by Val"
FT      Misc-difference 821
FT      /note= "Wild-type Asn substituted by Ser"
FT      Misc-difference 1052
FT      /note= "Wild-type Ala substituted by Val"

```

FT Misc-difference 1062 /note- "Wild-type Gly substituted by Ala"
FT Misc-difference 1118 /note- "Wild-type Lys substituted by Met"
FT Misc-difference 1151 /note- "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note- "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note- "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note- "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note- "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note- "Wild-type Ser substituted by Thr"
FT Misc-difference 1930 /note- "Wild-type Ser substituted by Thr"
PN US6423318-B1.
PD 23-JUL-2002.
31-AUG-2000; 2000US-0653499.
07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
DR WPI: 2002-680946/73.
DR N-PSDB; ABS52788.
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Example 3; Column 67-78; 71pp; English.
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDEETIAANSKD 20
DB 922 KVNPHGMLDEETIAANSKD 941
RESULT 11
ABG31729
ID ABG31729 standard; Protein: 2227 AA.
XX
XX ABG31729;
AC
DT 29-NOV-2002 (first entry)
XX
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
DE
XX
KM Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
KM HAV 4380.
XX
XX Hepatitis A virus strain HM-175.
OS
XX US6423318-B1.
PN

XX
PD 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-0653499.
PF
XX
PR 07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
DR WPI: 2002-680946/73.
DR N-PSDB; ABS52789.
XX
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Disclosure; Column 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents an attenuated
CC hepatitis A virus 4830 polypeptide.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPHGMLDEETIAANSKD 20
DB 922 KVNPHGMLDEETIAANSKD 941
RESULT 12
AAE19899
ID AAE19899 standard; Protein: 2227 AA.
XX
XX AAE19899;
AC
XX
DT 18-JUN-2002 (first entry)
XX
XX Hepatitis A virus (HAV) protein.
DE
XX
KM Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KM cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
OS
XX WO200213855-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 15-AUG-2001; 2001WO-1B01808.
PF
XX
PR 17-AUG-2000; 2000US-225767P.
PR 29-AUG-2000; 2000US-229175P.
PR 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPER AB.
PA
XX
XX Sallberg M, Hultgren C;
PI
XX
XX WPI: 2002-241837/29.
DR N-PSDB; AAD31766.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has

PT epitope present in hepatitis C virus -
XX
PS Claim 11; Page 82-87; 120pp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to
CC a hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KVNPFHGMLEIEIANSKD 20
|||||
922 KVNPFHGMLEIEIANSKD 941
DB
RESULT 13
ABU08639
ID ABU08639 standard; Protein; 2227 AA.
XX
AC ABU08639;
XX
DT 03-JUN-2003 (first entry)
XX
DE Wild type human hepatitis A virus strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection.
XX
OS Hepatitis A virus strain HM-175.
XX
PN US2002176869-A1.
XX
PD 28-NOV-2002.
XX
PF 29-APR-2002; 2002US-0135988.
XX
PR 07-JUN-1995; 95US-0475886.
PR 31-AUG-2000; 2000US-0653499.
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-0508610.
PR 17-APR-1995; 95US-0397232.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX
DR MPI: 2003-352605/33.
DR N-PSDB; ABX93473.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease -
XX
PS Disclosure; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175.
XX

SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KVNPFHGMLEIEIANSKD 20
|||||
922 KVNPFHGMLEIEIANSKD 941
DB
RESULT 14
ABU08640
ID ABU08640 standard; Protein; 2227 AA.
XX
AC ABU08640;
XX
DT 03-JUN-2003 (first entry)
XX
DE Attenuated (pass35) hepatitis A virus strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
OS Hepatitis A virus strain HM-175.
XX
PN US2002176869-A1.
XX
PD 28-NOV-2002.
XX
PF 29-APR-2002; 2002US-0135988.
XX
PR 07-JUN-1995; 95US-0475886.
PR 31-AUG-2000; 2000US-0653499.
PR 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-0508610.
PR 17-APR-1995; 95US-0397232.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX
DR MPI: 2003-352605/33.
DR N-PSDB; ABX93474.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease -
XX
PS Example 3; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 104; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KVNPFHGMLEIEIANSKD 20
|||||
922 KVNPFHGMLEIEIANSKD 941
DB
RESULT 15
ABU08641

ID AB008641 standard; Protein: 2227 AA.
 AC AB008641;
 XX
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DE Attenuated hepatitis A virus (4380) strain HM-175.
 XX
 KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
 XX vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
 OS Hepatitis A virus strain HM-175.
 PN US2002176869-A1.
 PD 28-NOV-2002.
 PD 29-APR-2002; 2002US-0135988.
 PF 07-JUN-1995; 95US-0475886.
 PR 31-AUG-2000; 2000US-0653499.
 PR 18-SEP-1992; 92US-0947338.
 PR 17-SEP-1993; 93WO-US08610.
 PR 17-APR-1995; 95US-0397232.
 XX
 PA (FUNK/) FUNKHOUSER A W.
 PA (EMER/) EMERSON S U.
 PA (PURC/) PURCELL R H.
 PA (DHON/) D'HONDT E.
 PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
 DR WPI: 2003-352605/33.
 DR N-PSDB: ABX93475.
 XX
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
 PT useful in vaccines for protecting primates against hepatitis infection
 PT and disease -
 XX
 PS Disclosure; Page 45-51; 70pp; English.
 XX
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
 CC a vaccine for protecting primates against hepatitis infection and
 CC disease. This is the amino acid sequence of an attenuated human
 CC hepatitis A virus (4380) strain HM-175.
 XX
 SO Sequence 2227 AA;
 XX
 XX Query Match 100.0%; Score 104; DB 24; Length 2227;
 XX Best Local Similarity 100.0%; Pred. No. 2.8e-08;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 KVNPPHGMIDLEETFAANSKD 20
 DB 922 KVNPPHGMIDLEETFAANSKD 941

Search completed: October 1, 2003, 09:56:43
 Job time : 44.5294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds
(without alignments) 149.988 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104
Sequence: 1 KVNPPHGMJDLIEIANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	2227	1 GNNYHM	genome polyprotein
2	104	100.0	2227	1 GNNYHR	genome polyprotein
3	104	100.0	2227	1 GNNYMK	genome polyprotein
4	104	100.0	2227	1 GNNYHB	genome polyprotein
5	101	97.1	2230	1 GNNYSA	genome polyprotein
6	47	45.2	333	2 A12485	hypothetical prote
7	47	45.2	387	2 B86669	RecA protein [Impo
8	47	45.2	506	1 S58522	glycine-tRNA 119as
9	47	45.2	1060	2 F88710	protein C01G5.4 [1
10	47	45.2	1079	2 T30996	hypothetical prote
11	46	44.2	259	2 F82087	deoxyribose-phosph
12	46	44.2	444	2 C89768	conserved hypotet
13	46	44.2	677	2 H64574	DNA topoisomerase
14	45	43.3	195	2 T08812	probable succinate
15	45	43.3	383	2 F96582	hypothetical prote
16	45	43.3	417	2 A44529	hypothetical prote
17	45	43.3	852	2 B72685	succinate-CoA liga
18	45	43.3	935	2 T19011	hypothetical prote
19	45	43.3	1213	2 T19835	hypothetical prote
20	44	42.3	223	2 T37962	hypothetical prote
21	44	42.3	344	2 D97761	hypothetical prote
22	44	42.3	500	2 S50508	hypothetical prote
23	43	41.8	373	2 A69773	ANP protein - yea
24	43	41.3	140	2 E81659	hypothetical prote
25	43	41.3	141	2 G71501	conserved hypotet
26	43	41.3	257	2 AF2592	hypothetical prote
27	43	41.3	309	2 H97374	2-deoxyribose-5-ph
28	43	41.3	359	2 H71089	2-deoxyribose-5-ph
29	43	41.3	327	2 T00876	hypothetical prote
					probable glycerate

30	43	41.3	440	2 B71858	adenylosuccinate 1
31	43	41.3	440	2 H64658	adenylosuccinate 1
32	43	41.3	609	2 AB0955	glutamine-fructose
33	43	41.3	917	2 T21870	hypothetical prote
34	42.5	40.9	599	2 A86810	1-deoxyxylulose-5-
35	42	40.4	222	2 C82343	conserved hypotet
36	42	40.4	229	2 S77449	hydrogenase expres
37	42	40.4	259	2 D91296	2-deoxyribose-5-ph
38	42	40.4	290	2 S19426	hypothetical prote
39	42	40.4	309	2 T33259	hypothetical prote
40	42	40.4	342	2 C72313	hypothetical prote
41	42	40.4	365	2 B48945	recombination prote
42	42	40.4	384	2 E82088	conserved hypotet
43	42	40.4	495	2 D64578	conserved hypotet
44	42	40.4	539	2 G70520	probable csp prote
45	42	40.4			

ALIGNMENTS

RESULT 1

GNNYHM genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:Chen, J.L.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with
A:Reference number: A25981. NCID:87061253; PMID:3023706
A:Accession: A25981

A:Molecule type: genomic RNA
A:Residues: 1-2227 <COR>
A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolyase; nucleotidyltr
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP4>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLIEIANSKD 20
DB 922 KVNPPHGMJDLIEIANSKD 941

RESULT 2

GNNYHR

genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAU>
A:Cross-references: GB:R02990; NID:g3329596; PIDN:AAA45471.1; PID:g3329597
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20
|||||
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 3
GNNYK
genome polypeptide - human hepatitis A virus (strain HM-175/TMK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g3329594; PIDN:AAA45471.1; PID:g3329595
A:Note: submitted to Genbank, August 1987
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20
|||||
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 4
GNNYB
-genome polypeptide - human hepatitis A virus (strain HMB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
Vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Klein, R.; Wimmer, E.; Delnhard
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP4>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein vrg #status predicted <VP6>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20
|||||
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 5
GNNYSA
genome polypeptide - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JFID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222597; PIDN:BA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA03490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardize, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; polypeptide
F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:350-495/Product: coat protein 1C #status predicted <C1C>
 F:496-795/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.1%; Score 101; DB 1; Length 2230;
 Best Local Similarity 95.0%; Pred. No. 4, 6e-08;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEIANSKD 20
 DB 926 KVNPHGMLDEIANSKD 945

RESULT 6

hypothetical protein all7065 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C:Accession: A12485
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kultz, T.; Sasamoto, S.; Matsumoto, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MID:21595285; PMID:11759840
 A:Status: preliminary
 A:Accession: A12485
 A:Molecule type: DNA
 A:Residues: 1-333 <KUR>
 A:Cross-references: GB:BA000020; PIDN:BA078149.1; PID:q12135603; GSPDB:GN00180
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all7065
 A:Genome: plasmid

Query Match 45.2%; Score 47; DB 2; Length 333;
 Best Local Similarity 40.0%; Pred. No. 5, 6;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEIANSKD 20
 57 KSGYPRGKITRDEVANVOD 76

RESULT 7

Reca protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: B86669
 R:Boletín, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weisenbach, J.; Ehrlich, S.
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.s.
 A:Reference number: AB6625; MID:21235186; PMID:11337471
 A:Accession: B86669
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-387 <STO>
 A:Cross-references: GB:AE005176; PID:q12723223; PIDN:AAK04452.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: reca
 C:Superfamily: recombination protein reca

Query Match 45.2%; Score 47; DB 2; Length 387;
 Best Local Similarity 45.0%; Pred. No. 6, 7;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 OY 1 KVNPHGMLDEIANSKD 20
 DB 340 KVRTAHGLDEAEVATETD 359

RESULT 8

glycine-tRNA ligase (EC 6.1.1.14) - Thermus aquaticus
 S58522
 N:Alternate names: glycyl-tRNA synthetase
 C:Species: Thermus aquaticus
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S58522
 R:Logan, D.T.; Mazauric, M.H.; Kern, D.; Moras, D.
 EMBO J. 14, 4156-4167, 1995
 A:Title: Crystal structure of glycyl-tRNA synthetase from Thermus thermophilus.
 A:Reference number: S58522; MID:96016187; PMID:7556056
 A:Accession: S58522
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-506 <LOG>
 A:Experimental source: strain HB8
 A:Note: The source is designated as Thermus thermophilus
 C:Superfamily: Mycoplasma genitalium glycine-tRNA ligase
 C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 45.2%; Score 47; DB 1; Length 506;
 Best Local Similarity 75.0%; Pred. No. 9, 2;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 FPHGMLDEIANSKD 15
 DB 299 FPHGMLDEIANSKD 310

RESULT 9

protein C01G5.4 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: F88710
 R:anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
 A:Reference number: A75000; MID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
 A:Accession: F88710
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1060 <STO>
 A:Cross-references: GB:chr_IV; PIDN:AA037736.1; PID:q1208856; GSPDB:GN00022; CESP:C01
 C:Genetics:
 A:Gene: C01G5.4
 A:Map position: 4

Query Match 45.2%; Score 47; DB 2; Length 1060;
 Best Local Similarity 36.8%; Pred. No. 22;
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNFPHGMLDEIANSKD 20
 DB 391 IGFSRGITLDKHYAGDARD 409

RESULT 10

hypothetical protein C01G5.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T30996
 R:Bradshaw, H.; Steillies, L.

submitted to the EMBL Data Library, August 1999
A:Description: The sequence of C. elegans cosmid C01G5.

A:Reference number: 220956

A:Accession: J30996

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1079

A:Cross-references: EMBL:U50068; PIDN:AB37736.2

A:Experimental source: Strain Bristol N2

C:Genetics:

A:Map position: IV

A:Introns: 10/1; 31/1; 75/1; 108/1; 144/1; 927/2; 1056/2

A:Note: C01G5.4

Query Match 45.2%; Score 47; DB 2; Length 1079;
Best Local Similarity 36.8%; Pred. No. 22;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEETIAANSKD 20

410 IGFSGRGLDKHVGADMD 428

RESULT 11

F82087

deoxyribose-phosphate aldolase VC2350 [imported] - Vibrio cholerae (strain N16961 serogr C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: F82087

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamatyan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: F82087

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-259 <HE>

A:Cross-references: GB:AE004305; GB:AE003852; NID:99656912; PIDN:AAF95493.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2350

A:Map position: 1

C:Superfamily: deoxyribose-phosphate aldolase

Query Match 44.2%; Score 46; DB 2; Length 259;
Best Local Similarity 52.9%; Pred. No. 6.2;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

3 NPPHGMIDLEETIAANSK 19

75 NPPHGMIDLEETIAVETK 91

Db

RESULT 12

C89768

conserved hypothetical protein SA0083 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: C89768

R:Kuroda, M.; Ohts, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud

ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: C89768

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-444 <RU>

A:Cross-references: GB:BA000018; PID:q13700003; PIDN:BA041302.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:
A:Gene: SA0083

Query Match 44.2%; Score 46; DB 2; Length 444;

Best Local Similarity 52.6%; Pred. No. 12;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEETIAANSKD 20

384 VNPPHGMIDLEETIAANSKD 402

Db

RESULT 13

H64574

DNA topoisomerase I - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999

C:Accession: H64574

R:Tombs, J.F.; White, O.; Kervat, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: H64574

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-677 <TOM>

A:Cross-references: GB:AE000559; GB:AE000511; NID:92313536; PIDN:AD07502.1; PID:9231

C:Superfamily: DNA topoisomerase I

Query Match 44.2%; Score 46; DB 2; Length 677;
Best Local Similarity 31.6%; Pred. No. 19;

Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 KVNPPHGMIDLEETIAANSK 19

346 RITHPHALKDEKVCSDAK 364

Db

RESULT 14

T08812

probable succinate-CoA ligase (GDP-forming) (EC 6.2.1.4) beta chain - human (fragment

N:Alternate names: protein DKFZp586M2023.1; succinyl-CoA synthetase (GDP-forming) bet

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002

C:Accession: T08812

R:Angorge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16472

A:Accession: T08812

A:Molecule type: mRNA

A:Residues: 1-195 <ANS>

A:Cross-references: EMBL:AL050226

A:Experimental source: adult uterus; clone DKFZp586M2023

C:Genetics:

A:Note: DKFZp586M2023.1

C:Function:

A:Description: catalyzes the formation of succinyl-CoA from succinate with concomitan

C:Superfamily: succinate-CoA ligase (ADP-forming) beta chain

C:Keywords: acid-thiol ligase; coenzyme A; mitochondrion

Query Match 43.3%; Score 45; DB 2; Length 195;
Best Local Similarity 42.1%; Pred. No. 6.6;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEETIAANSKD 20

65 VSGPQGVWDIEVAASNP 83

Db

RESULT 15

F96582
 Hypothetical protein F1511.25 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: F96582
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Deward, K.;
 ansen, N.F.; Hughes, B.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MID:21016719; PMID:11130712
 A:Accession: F96582
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <STO>
 A:Cross-references: GB:AE005173; NID:94587558; PIDN:AAD25789.1; GSPDB:GN00141
 A:Genetics:
 A:Gene: F1511.25
 A:Map position: 1

Query Match 43.3%; Score 45; DB 2; Length 383;
 Best Local Similarity 44.4%; Pred. No. 14;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 3 NPPHGMLDIEIRIANSKD 20
 : : 1 1 1 1 1 1 1 1
 Db 209 DYPVGLDIEKITLPPGKD 226

Search completed: October 1, 2003, 10:04:40
 Job time : 14.8235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds

(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNPHGMJDEEIANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_41:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	104	100.0	2226	1	POLG_HPAV4	P26581 hepatitis a
2	104	100.0	2226	1	POLG_HPAV8	P26582 hepatitis a
3	104	100.0	2227	1	POLG_HPAV8	P08617 hepatitis a
4	104	100.0	2227	1	POLG_HPAV4	P06441 hepatitis a
5	104	100.0	2227	1	POLG_HPAV4	P13901 hepatitis a
6	101	97.1	2230	1	POLG_HPAV5	P14553 simian hepa
7	100	96.2	2226	1	POLG_HPAV2	P26580 hepatitis a
8	100	95.2	387	1	REC2_LACUA	P01840 lactococcus
9	47	45.2	404	1	SCB2_MOUSE	P92218 mus musculu
10	47	45.2	505	1	SYG_THETH	P56206 thermus the
11	46	44.2	259	1	DEOC_VIBCH	P09617 vibrio chol
12	45	43.3	432	1	SCB2_HUMAN	P09619 homo sapien
13	45	43.3	433	1	SCB2_PIG	P53590 sus scrofa
14	44	42.3	500	1	ANP1_YEAST	P32629 saccharomy
15	44	41.3	98	1	SRL1_HUMAN	P07111 homo sapien
16	43	41.3	259	1	DEOC_NGR75	P08409 agrobacteri
17	43	41.3	357	1	UPB2_CHICK	P07429 gallus gall
18	43	41.3	440	1	PUR8_HELPY	P07429 gallus gall
19	43	41.3	440	1	PUR8_HELPY	P56468 helicobacte
20	43	41.3	608	1	GLMS_SALTI	P08292 s gluccosam
21	42	40.4	214	1	GLP1_PHOAT	P45853 pharbitis n
22	42	40.4	259	1	DEOC_ECOS7	P08236 escherichia
23	42	40.4	290	1	YC06_YEAST	P23617 saccharomy
24	41	39.9	406	1	CALB_PROSL	P08919 proteus sp.
25	41	39.4	193	1	YB71_HAEIN	P44339 haemophilus
26	41	39.4	218	1	Y363_RICPR	P09296 rickettsia
27	41	39.4	400	1	NIFS_ANAAT	P43884 anabaena az
28	41	39.4	400	1	NIFS_ANAAT	P12623 anabaena sp
29	41	39.4	502	1	NUZC_MESVI	P09496 mesostigma
30	41	39.4	552	1	SYO_CLOPE	P08493 clostridium
31	40.5	38.9	237	1	VIB8_AGR75	P17798 agrobacteri
32	40.5	38.9	434	1	UDG_RICPR	P05973 rickettsia
33	40.5	38.9	1286	1	RPOL_VACCC	P20504 vaccinia vi

34	40.5	38.9	1286	1	RPOL_VAAR	P33053 variola vir
35	40.5	38.9	1287	1	RPOL_VACCV	P07392 vaccinia vi
36	40	38.5	126	1	YF81_XYLF	P09612 xyella fas
37	40	38.5	310	1	YCCK_BACSU	P46905 bacillus su
38	40	38.5	353	1	UPB2_MOUSE	P08623 mus musculu
39	40	38.5	404	1	ISCS_NEIMA	P091x0 neisseria m
40	40	38.5	446	1	SYG_MYCGE	P47493 mycoplasma
41	40	38.5	507	1	YIR4_YEAST	P40486 saccharomy
42	40	38.5	558	1	PILE_NEIGO	P37094 neisseria g
43	40	38.5	595	1	KM2_HUMAN	P091r3 homo sapien
44	40	38.5	605	1	UPB2_HUMAN	P07504 homo sapien
45	40	38.5	646	1	KDBE_SCHPO	P10364 schizosacch

ALIGNMENTS

RESULT 1
POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 28-FEB-2003 (rel. 41, Last annotation update)
DE Genome polyprotein (contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)).
DE P3D (EC 2.7.7.48)).
OS Hepatitis A virus (strain 43C).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RT J. Virol. 65:2056-2065(1991).
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(n).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@sdb.ch).
CC
CC EMBL; M59809; AAA54569.1; -
CC MEROPS; C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVL.
DR Pfam; PF00680; RNA_dep_RNA_pol.1.
DR Pfam; PF00910; RNA_helicase.1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CORE PROTEIN VP4 (P1A).
FT COAT PROTEIN VP2 (P1B).
FT COAT PROTEIN VP3 (P1C).
FT COAT PROTEIN VP1 (P1D).
FT CORE PROTEIN P2A.
FT CORE PROTEIN P2B.
FT CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA: 251107 MW: 403B4CA80B9BF75 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDEETIANSKD 20
Db 922 KVNPHGMLDEETIANSKD 941
|||||

RESULT 2
POLG_HPAV8 STANDARD; PRT: 2226 AA.
ID POLG_HPAV8
AC P26582;
DE 01-AUG-1992 (Rel. 23, Created)
DE 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (RC 2.7.7.48)]
DE Hepatitis A virus (strain 18f)
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemmon S.M., Murphy P.C., Shields P.A., Ping L.H., Felstone S.M.,
RA Cromeans T., Jansen R.W.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RT J. Virol. 65:2056-2065(1991).
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M59808; AAA45467.1; -
CC PDB: 1Q47; 15-MAY-00.
CC MEROPS: C03.005.-
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC Polyprotein: Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
CC
CC CHAIN 1 23
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).
FT CHAIN 795 900 COAT PROTEIN VP1 (PID).
FT CHAIN 901 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA: 251292 MW: 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLDEETIANSKD 20
Db 922 KVNPHGMLDEETIANSKD 941
|||||

RESULT 3
POLG_HPAV8 STANDARD; PRT: 2227 AA.
ID POLG_HPAV8
AC P08617; P06443; Q81082;
DE 01-AUG-1988 (Rel. 08, Created)
DE 01-AUG-1988 (Rel. 08, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (RC 2.7.7.48)]
DE Hepatitis A virus (strain HM-175).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.,
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses."
RT J. Virol. 61:50-59(1987).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Felstone S.M.,
RA Purcell R.H.,
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus."
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RL [3]
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Matzel J.V. Jr.,
RA Purcell R.H., Felstone S.M.,
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase."
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF. 3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@sb-sib.ch).

CC EMBL: M14114; AAA45475.1; -
 DR EMBL: M14707; AAA45465.1; -
 DR EMBL: M14707; AAA45466.1; ALT_INIT.
 DR EMBL: M16632; AAA45471.1; -
 DR PIR: A25981; GNNYHM.
 DR PIR: A94149; GNNYHM.
 DR PDB: 1HAV; 23-DEC-96.
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calic1.pol_hel.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_PSVlr.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 KM Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.

FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1062 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 FT SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPFHGMGLDELIANSKD 20
 922 KVNPFHGMGLDELIANSKD 941

RESULT 4
 POLG_HPAVL STANDARD; PRT: 2227 AA.
 ID POLG_HPAVL
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain LA).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., van Nest G., Dina D.;

RT "Primary structure and gene organization of human hepatitis A virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC [RNA](N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sb-sib.ch).

CC EMBL: K02990; AAA45472.1; -
 DR PIR: A03903; GNNYHM.
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calic1.pol_hel.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_PSVlr.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 KM Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1076
 FT CHAIN 1077 1422
 FT CHAIN 1423 1484
 FT CHAIN 1485 1507
 FT CHAIN 1508 1678
 FT CHAIN 1679 2227
 FT SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPFHGMGLDELIANSKD 20
 922 KVNPFHGMGLDELIANSKD 941

RESULT 5
 POLG_HPAVM STANDARD; PRT: 2227 AA.
 ID POLG_HPAVM
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain MEB).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E., Delhardt F.;

```
RT "The entire nucleotide sequence of the genome of human hepatitis A
RL virus (isolate MB).";
RT Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INJECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: M20273; AAA5474.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2; je-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLEETANSKD 20
Db ||||||||||||||||
922 KVNPHGMLEETANSKD 941

RESULT 6
POLG_HPVS STANDARD; PRT: 2230 AA.
AC P14533;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
```

```
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapartidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: D00924; BAA00766.1; -.
DR EMBL: X15461; CA33490.1; -.
DR PIR: A30470; GNNVSA.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calicet_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICIVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 97.1%; Score 101; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 7; 7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLEETANSKD 20
Db ||||||||||||||||
926 KVNPHGMLEETANSKD 945

RESULT 7
POLG_HPAV2 STANDARD; PRT: 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
```

OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12094;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-91162758; PubMed-1705995;
RA Lemon S.M., Murphy P.C., Shelds P.A., Ping L.H., Feinstone S.M.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PFM: SPECIFIC ENZYMAIC CLEAVAGES IN VITRO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M59810; AAA45468.1; -
DR MEROPS: C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICVIRUSNS.
KM Polyprotein; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SO SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
Query Match 96.2%; Score 100; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1,1e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNPFHGMLEIEIANSKD 20
DB 922 KVNPFHGMLEIEIANSKE 941
ID RECD_LACLA STANDARD: PRT: 387 AA.
AC 001840; 09CIR3;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RecA protein, chromosomal (Recombinase A).
GN RECA OR L10354.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI_TaxID-1360;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-ML3;
RX MEDLINE-92384590; PubMed-1514816;
RA Duval P., Ehrlich S.D., Gruss A.;
RT "Use of degenerate primers for polymerase chain reaction cloning and
RT sequencing of the Lactococcus lactis subsp. lactis recA gene."
RL Appl. Environ. Microbiol. 58:2674-2678(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-IL1403;
RX MEDLINE-2135186; PubMed-11337471;
RA Bolotin A., Winkler P., Manger S., Jallion O., Malarne K.,
RA Weisenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis 111403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF
CC SINGLE-STRANDED DNA, THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED
CC DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF
CC HOMOLOGOUS SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING
CC ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the recA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M88106; AAA25216.1; -
DR EMBL: AE006272; AAK04452.1; -
DR PIR: B86669; B86669.
DR HAMAP: P26345; 1G19.
DR InterPro: IPR003593; AAA_Arpase.
DR InterPro: IPR001553; RecA.
DR Pfam: PF00154; recA.1.
DR PRINTS: PR00142; RECA.
DR ProDom: PD000229; RECA.1.
DR SMART: SM00382; AAA.1.
DR PROSITE: PS00321; RECA_1.1.
DR PROSITE: PS50162; RECA_2.1.
DR PROSITE: PS50163; RECA_3.1.
KM DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;
KM Complete proteome.
FT NP_BIND 80 87 ATP (BY SIMILARITY).
FT CONFLICT 288 288 R -> S (IN REF. 1).
FT CONFLICT 288 288 D -> E (IN REF. 1).
FT CONFLICT 332 333 E -> DV (IN REF. 1).
FT CONFLICT 343 343 T -> A (IN REF. 1).
FT CONFLICT 350 351 EA -> DS (IN REF. 1).
FT CONFLICT 357 387 TSDSTKATATATKAEKVEVTEIELELED -> EETETA
FT FKN (IN REF. 1).
SO SEQUENCE 387 AA; 41477 MW; A8CAFC0B827BF14F CRC64;
Query Match 45.2%; Score 47; DB 1; Length 387;
Best Local Similarity 45.0%; Pred. No. 2.3;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 KVNPFHGMLEIEIANSKD 20
DB 340 KVRTAHGLDEVAETTED 359
ID SCB2_MOUSE STANDARD: PRT: 404 AA.
AC 092218;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RecA protein, chromosomal (Recombinase A).
GN RECA OR L10354.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Succinyl-CoA ligase [GDP-forming], beta-chain, mitochondrial precursor
 DE (EC 6.2.1.4) (Succinyl-CoA synthetase, beta chain) (SCS-beta) (GTP-
 specific succinyl-CoA synthetase beta subunit) (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart.
 RX MEDLINE=98438536; PubMed=9765291;
 RA Johnson J.D., Mehus J.G., Tews K., Milavetz B.I., Lambeth D.O.;
 RT "Genetic evidence for the expression of ATP- and GTP-specific
 RT succinyl-CoA synthetases in multicellular eucaryotes.";
 RL J. Biol. Chem. 273:27580-27586(1998).
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA -> GDP + succinyl-CoA +
 CC phosphate.
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
 CC ACID CYCLE.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
 CC subunit family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF058956; AAC64399.1; -
 DR HSSP: P07460; 1SCU
 DR MGD; MGI:1306824; Suc1g2.
 DR InterPro: IPR003135; ATP-grasp.
 DR InterPro: IPR005809; CoA_lig_beta.
 DR InterPro: IPR005811; CoA_ligase.
 DR Pfam: PF02222; ATP-grasp.1.
 DR Pfam: PF00549; ligase-CoA.1.
 DR TIGRPFAMs: TIGR01016; succoAdelta.1.
 DR PROSITE; PS01217; SUCCINYL-CoA_lig_3; 1.
 DR Ligase; Glycyllys; Tricarboxylic acid cycle; Mitochondrion;
 KW TRANSIT PEPTIDE.
 KW NON TER
 FT CHAIN
 FT 10 404 SUCCINYL-CoA LIGASE [GDP-FORMING] BETA-
 FT CHAIN.
 SQ SEQUENCE 404 AA; 43857 MW; 511757A75883E8C4 CRC64;
 FT
 QY 2 VFNPQGLDELEIANSKD 20
 Db 140 VGSQGRSDIEVVAASPE 158
 FT
 RESULT 10
 SYG_THETH
 ID SYG_THETH STANDARD; PRT; 505 AA.
 AC P56206; O50551;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glycyl-tRNA synthetase (EC 6.1.1.14) (glycine--tRNA ligase) (GLYRS).
 GN GLYS.
 OS Thermus thermophilus.

CC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 CC Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.; AND X-RAY CRYSTALLOGRAPHY (2.75 ANGSTROMS).
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=96016187; PubMed=756056;
 RA Logan D.T., Mazauric M.-H., Kern D., Moras D.;
 RT "Crystal structure of glycyl-tRNA synthetase from Thermus
 RT thermophilus.";
 RL EMBO J. 14:4156-4167(1995).
 CC [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=98149692; PubMed=9490048;
 RA Mazauric M.-H., Keith G., Logan D., Kreutzler R., Giege R., Kern D.;
 RT "Glycyl-tRNA synthetase from Thermus thermophilus - wide structural
 RT divergence with other prokaryotic glycyl-tRNA synthetases and
 RT functional inter-relation with prokaryotic and eukaryotic glycylation
 RT systems.";
 RL Eur. J. Biochem. 251:744-757(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) -> AMP + diphosphate
 CC + glycyl-tRNA(Gly).
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AJ222643; CAA10903.1; -
 DR PDB: 1AT1; 07-JUL-97.
 DR PDB: 1B76; 28-JAN-99.
 DR PDB: 1GGM; 28-JAN-99.
 DR HAMAP; MF_00253; -; 1.
 DR InterPro: IPR004154; HGTP_anticodon.
 DR InterPro: IPR002314; tRNA-synt_2b.
 DR InterPro: IPR002315; tRNA-synt_gly.
 DR InterPro: IPR006195; tRNA_ligase_lig.
 DR Pfam; PF03129; HGTP_anticodon.1.
 DR Pfam; PF00587; tRNA-synt_2b; 1.
 DR PRINTS; PRO1043; TRNASYNTGLY.
 DR TIGRPFAMs: TIGR00389; glys_dimeric.1.
 DR PROSITE; PS50862; AA_tRNA_LIGASE_II; 1.
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW 3D-structure.
 FT INIT MET
 FT CONFLICT 1 0
 FT CONFLICT 115 119 YRAME -> TEPAR (IN REF. 2).
 FT CONFLICT 191 199 QGIVNFRN -> RASSSTST (IN REF. 2).
 FT CONFLICT 215 219 IGKAF -> SARPS (IN REF. 2).
 FT CONFLICT 266 266 E -> R (IN REF. 2).
 FT CONFLICT 283 284 SS -> EL (IN REF. 2).
 FT CONFLICT 302 303 SL -> LE (IN REF. 2).
 FT CONFLICT 310 310 Q -> N (IN REF. 2).
 FT HELIX 5 14
 FT TURN 15 16
 FT STRAND 18 20
 FT TURN 21 22
 FT HELIX 23 25
 FT TURN 26 26
 FT TURN 29 30
 FT STRAND 32 34
 FT TURN 36 53
 FT HELIX 54 55
 FT STRAND 60 64
 FT STRAND 68 70
 FT TURN 71 72

```

FT HELIX 73 76
FT TURN 77 77
FT HELIX 79 82
FT TURN 83 83
FT STRAND 168 168
FT STRAND 170 172
FT HELIX 179 181
FT STRAND 182 185
FT HELIX 190 195
FT TURN 196 196
FT HELIX 197 204
FT TURN 205 205
FT STRAND 210 220
FT TURN 227 227
FT HELIX 228 230
FT STRAND 234 244
FT HELIX 246 248
FT HELIX 249 266
FT TURN 267 268
FT HELIX 271 273
FT STRAND 274 278
FT TURN 281 283
FT TURN 286 287
FT STRAND 290 298
FT TURN 299 300
FT STRAND 301 310
FT TURN 312 313
FT HELIX 314 319
FT TURN 323 327
FT STRAND 344 345
FT STRAND 352 353
FT STRAND 356 363
FT HELIX 364 375
FT STRAND 376 380
FT TURN 382 383
FT STRAND 386 390
FT HELIX 394 396
FT STRAND 401 405
FT HELIX 411 425
FT TURN 426 427
FT STRAND 432 434
FT HELIX 440 449
FT TURN 450 451
FT STRAND 454 458
FT HELIX 460 463
FT TURN 464 468
FT TURN 467 475
FT TURN 472 475
FT TURN 476 481
FT STRAND 482 484
FT STRAND 487 491
FT HELIX 492 503
SQ SEQUENCE 505 AA; 58081 MM; EC1C8D5386AB7570 CRC64;

Query Match 45.2%; Score 47; DB 1; Length 505;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 FPHGMLDEEIA 15
DB 298 FPHGSLDEGIA 309

RESULT 11
DEOC_VIRCH STANDARD: PRT: 259 AA.
AC 09FVL7-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyribosomalase)
DE (Deoxyribosomalase) (DERA).
GN DEOC OR VC2330.

```

```

OS Vibrio Cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_Taxid=666;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-El Tor N16961 / Serotype O1;
RX MEDLINE-20406833; PubMed-10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Hart D.H., Hickey E.R., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Baas S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT Cholerae."
RL Nature 406:477-483(2000).
CC -I- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate - D-
CC glycerinaldehyde 3-phosphate + acetaldehyde.
CC -I- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -I- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES. DEOC
CC SUBFAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF004305; AAP95493.1; .
DR PIR: F82087; F82087.
DR TIGR: VC2350; .
DR HAMAP: MF_00592; .; 1.
DR InterPro: IPR002915; DEOC.
DR Pfam: PF01791; DEOC; 1.
KW Lyase; Schiff base; Complete proteome.
FT BINDING 166 166 SCHIFF-BASE WITH ACETALDEHYDE (BY
FT ACT_SITE 200 200 SIMILARITY).
FT BY SIMILARITY.
SQ SEQUENCE 259 AA; 27959 MM; 287607AC6527C73 CRC64;

Query Match 44.2%; Score 46; DB 1; Length 259;
Best Local Similarity 52.9%; Pred. No. 2.2;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY 3 NFPHGMLDEEIAANSK 19
DB 75 NFPHGNDIEIAVAETK 91

RESULT 12
SCB2_HUMAN STANDARD: PRT: 432 AA.
ID SCB2_HUMAN 095195; O8W01;
AC 095199; 095195; O8W01;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial precursor
DE (EC 6.2.1.4) (Succinyl-CoA synthetase, beta-chain) (SCS-beta) (GTP-
DE specific succinyl-CoA synthetase beta subunit) (Fragment).
GN SUCUG2.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Uterus;
RC MEDLINE-22388257; PubMed-12477932;

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skala U., Smillius D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.",
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 [2]
 RC SEQUENCE OF 29-432 FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=98438536; Pubmed=9765291;
 RA Johnson J.D., Mehus J.G., Tews K., Mlavetz B.I., Lambeth D.O.,
 RT "genetic evidence for the expression of ATP- and GTP-specific
 RT succinyl-CoA synthetases in multicellular eucaryotes.",
 RL J. Biol. Chem. 273:27580-27586(1998).
 RP SEQUENCE OF 234-432 FROM N.A.
 RA Mel G., Yu W., Gibbs R.A.,
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA -> GDP + succinyl-CoA +
 CC phosphate.
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
 CC ACID CYCLE.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
 CC subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: BC007716; AA07716.1; -
 DR EMBL: BC019868; AA19868.1; -
 DR EMBL: AF038954; AAC6497.1; -
 DR EMBL: AF131748; AAD20032.1; -
 DR HSSP: P07460; 1SCU.
 DR Genew: HGNC:11450; SUCLG2.
 DR MIM: 603922; -
 DR GO: GO:0008325; C:succinate-CoA ligase complex (GDP-forming) . . .; NAS.
 DR GO: GO:0004776; F:succinate-CoA ligase (GDP-forming) activity; NAS.
 DR GO: GO:0006104; P:succinyl-CoA metabolism; NAS.
 DR InterPro: IPR003135; ATP-grasp.
 DR InterPro: IPR005809; CoA_lig_beta.
 DR InterPro: IPR005811; CoA_ligase.
 DR Pfam: PF02222; ATP-grasp.1.
 DR Pfam: PF00349; ligase-CoA.1.
 DR TIGRfam: TIGR01016; succoabeta.1.
 DR PROSITE: PS01217; SUCCINYL_COA_LIG_3; 1.
 DR Ligase: glycolysis; Tricarboxylic acid cycle; Mitochondrion;
 KM Transit peptide.
 FT NON_TER 1 1
 FT TRANSIT 1 37 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 38 432 SUCCINYL-CoA LIGASE [GDP-FORMING] BETA-
 FT CHAIN.
 FT CONFLICT 174 175 GV -> RS (IN REF. 2).

SQ SEQUENCE 432 AA; 46535 MW; EB4D025B42EA7BE0 CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 432;
 Best Local Similarity 42.1%; Pred. No. 5.6;
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 Oy 2 VNFPHGLDLEETANSKD 20
 Db 168 VGSFGGVDEIEVAASNPE 186
 RESULT 13
 ID SCB2_PIG STANDARD; PRT; 433 AA.
 AC P53590; Q95279;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial precursor
 DE (NC 6.2.1.4) (Succinyl-CoA synthetase, beta chain) (SCS-beta) (GTP-
 DE specific succinyl-CoA synthetase beta subunit) (Fragment).
 GN SUCLG2.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 ON NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-41.
 RC TISSUE=Heart;
 RX MEDLINE=94004462; Pubmed=8401211;
 RA Bailey D.L., Wolodko W.T., Bridger W.A.,
 RT "Cloning, characterization, and expression of the beta subunit of pig
 RT heart succinyl-CoA synthetase.",
 RL Protein Sci. 2:1255-1262(1993).
 RN [2]
 RP SEQUENCE OF 1-82 FROM N.A.
 RC TISSUE=Small Intestine;
 RX MEDLINE=96327607; Pubmed=8672129;
 RA Winteroe A.K., Fredholm M., Davies W.,
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library: analysis of 839 clones.",
 RL Mamm. Genome 7:509-517(1996).
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA -> GDP + succinyl-CoA +
 CC phosphate.
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
 CC ACID CYCLE.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta
 CC subunit family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL: L06944; AAA1120.1; ALT_INIT.
 DR EMBL: Z81187; CAB03559.1; -
 DR PDB: 1EUC; 27-JUL-00.
 DR PDB: 1EUD; 27-JUL-00.
 DR InterPro: IPR003135; ATP-grasp.
 DR InterPro: IPR005809; CoA_lig_beta.
 DR InterPro: IPR005811; CoA_ligase.
 DR Pfam: PF02222; ATP-grasp.1.
 DR Pfam: PF00349; ligase-CoA.1.
 DR TIGRfam: TIGR01016; succoabeta.1.
 DR PROSITE: PS01217; SUCCINYL_COA_LIG_3; 1.
 DR Ligase: glycolysis; Tricarboxylic acid cycle; Mitochondrion;
 KM Transit peptide; 3D-structure.
 FT NON_TER 1 1

FT TRANSIT <1 38 MITOCHONDRION.
 FT CHAIN 39 433 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-
 FT CHAIN.
 SQ SEQUENCE 433 AA; 46803 MW; AA04B72BC1B80E24 CRC64;
 Query Match 43.3%; Score 45; DB 1; Length 433;
 Best Local Similarity 42.1%; Pred. No. 5.6;
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 QY 2 VNFPHGMDLEIEAANSKD 20
 DB 169 VGSFGVDIEVMAASNP 187
 RESULT 14
 ANP1_YEAST STANDARD; PRT; 500 AA.
 AC P32629;
 DT 01-OCT-1993 (Rel. 27, Created)
 01-FEB-1995 (Rel. 31, Last sequence update)
 15-SEP-2003 (Rel. 42, Last annotation update)
 Mannan polymerase II complex ANP1 subunit (M-Pol II subunit ANP1)
 DE (Amanitophenyl propanediol resistance protein).
 GN ANP1 OR GEM3 OR YEL036C OR SYCP-ORF28.
 OS Saccharomyces cerevisiae (baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B-6441;
 RX MEDLINE=94016558; PubMed=8411151;
 RA "Mednick L., Sherman F.;
 "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 of Saccharomyces cerevisiae share a common ancestry.";
 RL J. Mol. Biol. 233:372-388(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972;
 RX MEDLINE=96096381; PubMed=9434768;
 RA Hashimoto H., Yoda K.;
 "Novel membrane protein complexes for protein glycosylation in the
 yeast Golgi apparatus.";
 RL Biochem. Biophys. Res. Commun. 241:682-686(1997).
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE=95045382; PubMed=7957057;
 RA Chapman R.E., Munro S.;
 "The functioning of the yeast Golgi apparatus requires an ER protein
 encoded by ANP1, a member of a new family of genes affecting the
 secretory pathway.";
 RL EMBO J. 13:4896-4907(1994).
 RN [5]
 RP ACTIVITY OF M-POL II COMPLEX, SUBUNITS, AND SUBCELLULAR LOCATION.
 RX MEDLINE=98094364; PubMed=9430634;
 RA Jungmann J., Munro S.;
 "Multi-protein complexes in the cis Golgi of Saccharomyces cerevisiae
 with alpha-1,6-mannosyltransferase activity.";
 RL EMBO J. 17:423-434(1998).
 RN [6]
 RP SUBCELLULAR LOCATION.

RX MEDLINE=20558560; PubMed=11095735;
 RA Todorow Z., Spang A., Carmack E., Yates J., Schekman R.;
 "Active recycling of yeast Golgi mannosyltransferase complexes through
 the endoplasmic reticulum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:13643-13648(2000)
 CC -1- FUNCTION: INVOLVED IN THE ORGANIZATION OF THE SECRETORY PATHWAY.
 CC REQUIRED TO MAINTAIN A FUNCTIONAL GOLGI APPARATUS.
 CC -1- FUNCTION: The M-Pol II complex possesses alpha-1,6-
 mannosyltransferase activity and is probably involved in the
 elongation of the mannan backbone of N-linked glycans on cell wall
 CC and periplasmic proteins.
 CC -1- SUBUNIT: Component of the M-Pol II complex composed of ANP1, MNN9,
 MNN10, MNN11 and HOC1.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Cis-Golgi.
 CC Recycles between endoplasmic reticulum and Golgi.
 CC -1- SIMILARITY: BELONGS TO THE ANP1 / MNN9 / VANI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; S65964; AAD13971.1; -;
 DR EMBL; I22171; AAA34426.1; -;
 DR EMBL; S66114; AAB28440.1; -;
 DR EMBL; I22173; AAA34937.1; -;
 DR EMBL; U18779; AAB65006.1; -;
 DR FIR; S50508; S50508.
 DR FIR; S5000762; ANP1.
 DR GO; GO:0000136; C:mannosyltransferase complex; TAS.
 DR GO; GO:0000009; F:alpha-1,6-mannosyltransferase activity; IDA.
 DR GO; GO:0000032; P:cell wall mannoprotein biosynthesis; TAS.
 DR GO; GO:0006487; P:N-linked glycosylation; IDA.
 DR InterPro; IPR005109; ANP1.
 DR Pfam; PF03452; ANP1; 1.
 KW Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;
 KW Endoplasmic reticulum.
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 16 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 28 500 LUMENAL (POTENTIAL).
 FT DOMAIN 446 473 GLN-RICH.
 FT CONFLICT 220 224 HHDKD -> QSGQGN (IN REF. 1).
 FT CONFLICT 313 313 F -> L (IN REF. 1).
 FT CONFLICT 472 500 PCKRPIDNDKKKKHKKRPKPEVLPDPPDRN -> RGNGLMT
 FT TTRTRKNITKKH (IN REF. 1).
 SQ SEQUENCE 500 AA; 58182 MW; 845B395CE54BCD14 CRC64;
 Query Match 42.3%; Score 44; DB 1; Length 500;
 Best Local Similarity 31.6%; Pred. No. 9.7;
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
 QY 2 VNFPHGMDLEIEAANSKD 20
 DB 106 MTPPHNLIDSLFVSDSSD 124
 RESULT 15
 SRG1_HUMAN STANDARD; PRT; 98 AA.
 AC 075711;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Scrape-responsive protein 1 precursor (SCRG-1).
 GN SCRG1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNPPHGMLEIRANSKD 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Minimum number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:**
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_proteint:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	251	12	Q9ENQ4
2	104	100.0	251	12	Q9ENQ7
3	104	100.0	251	12	Q9ENQ2
4	104	100.0	251	12	Q9ENP2
5	104	100.0	251	12	Q9ENQ6
6	104	100.0	251	12	Q9ENR1
7	104	100.0	251	12	Q9ENP1
8	104	100.0	251	12	Q9ENQ1
9	104	100.0	251	12	Q9ENQ9
10	104	100.0	251	12	Q9ENQ4
11	104	100.0	251	12	Q9ENP5
12	104	100.0	251	12	Q9ENP7
13	104	100.0	251	12	Q9ENP5
14	104	100.0	251	12	Q9ENP9
15	104	100.0	251	12	Q9ENP6
16	104	100.0	251	12	Q9ENN5

17	104	100.0	251	12	Q9ENN7	Q9enn7 hepatitis a
18	104	100.0	1124	12	Q84780	Q84780 hepatitis a
19	104	100.0	1161	12	Q05794	Q05794 hepatitis a
20	104	100.0	2216	12	Q9WMA2	Q9wma2 hepatitis a
21	104	100.0	2218	12	Q67824	Q67824 hepatitis a
22	104	100.0	2225	12	Q9DLJ2	Q9dlj2 hepatitis a
23	104	100.0	2227	12	Q9WMA3	Q9wma3 hepatitis a
24	104	100.0	2227	12	Q9WMA0	Q9wma0 hepatitis a
25	104	100.0	2227	12	Q67825	Q67825 hepatitis a
26	104	100.0	2227	12	Q9WMA1	Q9wma1 hepatitis a
27	104	100.0	2227	12	Q67826	Q67826 hepatitis a
28	104	100.0	2227	12	Q8VON6	Q8von6 hepatitis a
29	104	100.0	2227	12	Q9IFH5	Q9ifh5 hepatitis a
30	104	100.0	2227	12	Q9WMA4	Q9wma4 hepatitis a
31	98	94.2	251	12	Q9ENN3	Q9enn3 hepatitis a
32	98	94.2	251	12	Q9ENN0	Q9enn0 hepatitis a
33	98	94.2	251	12	Q9ENR0	Q9enr0 hepatitis a
34	98	94.2	251	12	Q9ENQ3	Q9enq3 hepatitis a
35	98	94.2	251	12	Q9ENP6	Q9enp6 hepatitis a
36	98	94.2	251	12	Q9ENQ8	Q9enq8 hepatitis a
37	98	94.2	251	12	Q9ENP3	Q9enp3 hepatitis a
38	98	94.2	251	12	Q9ENN8	Q9enn8 hepatitis a
39	98	94.2	251	12	Q9ENN0	Q9enn0 hepatitis a
40	98	94.2	251	12	Q9ENQ0	Q9enq0 hepatitis a
41	98	94.2	2227	12	Q8QV03	Q8qv03 hepatitis a
42	97	93.3	184	12	Q87092	Q87092 simian hepa
43	97	93.3	2225	12	Q9DMR1	Q9dmr1 hepatitis a
44	95	91.3	2218	12	Q67817	Q67817 hepatitis a
45	94	90.4	251	12	Q9ENQ2	Q9enq2 hepatitis a

ALIGNMENTS

RESULT 1
Q9ENQ4 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ4: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)
OS Polyprotein (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A201;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; .
FT NON_TER
FT NON_TER
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPPHGMLEIRANSKD 20
DB 86 KVNPPHGMLEIRANSKD 105
|||||

RESULT 2

Q9ENQ7 PRELIMINARY; PRT; 251 AA..
AC Q9ENQ7: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMblrel. 16, Last annotation update)

```
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A161;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047656; BAB12164.1; -.
FT NON_TER 1
FT SEQUENCE 251 AA; 28766 MW; A0EC02E3609C4D5B CRC64;

Query Match
Best Local Similarity 100.0%; Score 104; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPPHGMLDLEETANSKD 20
|||||
86 KVNPPHGMLDLEETANSKD 105

RESULT 3
Q9ENN2 PRELIMINARY; PRT; 251 AA.
AC Q9ENN2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A9;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047681; BAB12189.1; -.
FT NON_TER 1
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 104; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPPHGMLDLEETANSKD 20
|||||
86 KVNPPHGMLDLEETANSKD 105

RESULT 4
Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC Q9ENP2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujiwara K.;
```

```
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1
FT SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match
Best Local Similarity 100.0%; Score 104; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPPHGMLDLEETANSKD 20
|||||
86 KVNPPHGMLDLEETANSKD 105

RESULT 5
Q9EN06 PRELIMINARY; PRT; 251 AA.
AC Q9EN06;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A162;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1
FT SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;

Query Match
Best Local Similarity 100.0%; Score 104; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPPHGMLDLEETANSKD 20
|||||
86 KVNPPHGMLDLEETANSKD 105

RESULT 6
Q9ENR1 PRELIMINARY; PRT; 251 AA.
AC Q9ENR1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A1;
RA Fujiwara K.;
RT "Hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -.
FT NON_TER 1
FT SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match
Best Local Similarity 100.0%; Score 104; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMIDLEIANSKD 20
|||||

DB 86 KVNPHGMIDLEIANSKD 105

RESULT 7

ID O9ENP1 PRELIMINARY; PRT; 251 AA.

AC O9ENP1: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE Polypeptide (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

NC NCB1_TaxID=12092;

OX [1]

RA Fujikawa K.;

RT "hepatitis A virus.";

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB047672; BAB12180.1; -.

FT NON_TER 1 251

SQ SEQUENCE 251 AA; 28663 MW; C7EA66BD19A1619 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMIDLEIANSKD 20
|||||

DB 86 KVNPHGMIDLEIANSKD 105

RESULT 8

ID O9EN01 PRELIMINARY; PRT; 251 AA.

AC O9EN01: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE Polypeptide (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

NC NCB1_TaxID=12092;

OX [1]

RA Fujikawa K.;

RT "hepatitis A virus.";

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB047662; BAB12170.1; -.

FT NON_TER 1 251

SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMIDLEIANSKD 20
|||||

DB 86 KVNPHGMIDLEIANSKD 105

RESULT 9

O9EN09

ID O9EN09 PRELIMINARY; PRT; 251 AA.

AC O9EN09: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE Polypeptide (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

NC NCB1_TaxID=12092;

OX [1]

RA Fujikawa K.;

RT "hepatitis A virus.";

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB047654; BAB12162.1; -.

FT NON_TER 1 251

SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMIDLEIANSKD 20
|||||

DB 86 KVNPHGMIDLEIANSKD 105

RESULT 10

ID O9ENN4 PRELIMINARY; PRT; 251 AA.

AC O9ENN4: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE Polypeptide (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

NC NCB1_TaxID=12092;

OX [1]

RA Fujikawa K.;

RT "hepatitis A virus.";

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB047679; BAB12187.1; -.

FT NON_TER 1 251

SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 2.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMIDLEIANSKD 20
|||||

DB 86 KVNPHGMIDLEIANSKD 105

RESULT 11

ID O9ENP5 PRELIMINARY; PRT; 251 AA.

AC O9ENP5: 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DE Polypeptide (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

```

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RL "hepatitis A virus.";
DR Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF19C757A6D CRC64;

QY Query Match 100.0%; Score 104; DB 12; Length 251;
   Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QB 1 KVNPFGMLDEETIAANSKD 20
    |||||||
QB QENP7 PRELIMINARY; PRT; 251 AA.
AC QENP7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OC Hepatitis A virus.
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RL "hepatitis A virus.";
DR EMBL; AB047658; BABI216.1; -.
FT SUBMITTED (AUG-2000) TO THE EMBL/GENBANK/DDBJ DATABASES.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
   Best Local Similarity 100.0%; Pred. No. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QB 1 KVNPFGMLDEETIAANSKD 20
    |||||||
QB QENP7 PRELIMINARY; PRT; 251 AA.
AC QENP7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OC Hepatitis A virus.
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RL "hepatitis A virus.";
DR EMBL; AB047658; BABI216.1; -.
FT SUBMITTED (AUG-2000) TO THE EMBL/GENBANK/DDBJ DATABASES.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

```

FT	NON_TER	251	251		
SO	SEQUENCE	251 AA;	28720 MW;	C334248288219CA	CRC64;
Query Match					
Best Local Similarity		100.0%;		Score 104;	DB 12;
Matches		20; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 KVNFPHGMLDLEETANSKD 20				
DB	86 KVNFPHGMLDLEETANSKD 105				
RESULT 14					
OGENP9	PRELIMINARY;	PRT;	251 AA.		
AC	OGENP9;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment).				
OS	Hepatitis A virus.				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OC	Hepatovirus.				
NCBI_Taxid=12092;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-A303;				
RA	Fujiwara K.;				
RT	"hepatitis A virus.";				
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AB047664; BAB12172.1; -.				
FT	NON_TER	1	251		
FT	NON_TER	251	251		
SO	SEQUENCE	251 AA;	28752 MW;	7215A28AD2CA5C1A	CRC64;
Query Match					
Best Local Similarity		100.0%;		Score 104;	DB 12;
Matches		20; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 KVNFPHGMLDLEETANSKD 20				
DB	86 KVNFPHGMLDLEETANSKD 105				
RESULT 15					
OGENN6	PRELIMINARY;	PRT;	251 AA.		
ID	OGENN6				
AC	OGENN6;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment).				
OS	Hepatitis A virus.				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OC	Hepatovirus.				
NCBI_Taxid=12092;					
RP	SEQUENCE FROM N.A.				
RC	STRAIN-A713;				
RA	Fujiwara K.;				
RT	"hepatitis A virus.";				
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AB047677; BAB12185.1; -.				
FT	NON_TER	1	251		
FT	NON_TER	251	251		
SO	SEQUENCE	251 AA;	28720 MW;	C334248288219CA	CRC64;
Query Match					
Best Local Similarity		100.0%;		Score 104;	DB 12;
Matches		20; Conservative		0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 KVNFPHGMLDLEETANSKD 20				
DB	86 KVNFPHGMLDLEETANSKD 105				

Thu Oct 2 09:13:58 2003

us-09-171-432a-44.rpt

Page 5

Db 86 KVNPRGMIDLEIANSKD 105

Search completed: October 1, 2003, 10:02:40
Job time : 35.1765.secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 : Search time 12.7059 seconds
(without alignments)
66.600 Million cell updates/sec

Title: US-09-171-432a-44
Perfect score: 104
Sequence: 1 KVNFGMDLEIANSKD 20

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Dackfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	US-08-475-886-2	Sequence 2, Appl
2	104	100.0	2227	US-08-475-886-4	Sequence 4, Appl
3	104	100.0	2227	US-08-475-886-6	Sequence 6, Appl
4	104	100.0	2227	US-08-397-232-2	Sequence 2, Appl
5	104	100.0	2227	US-08-397-232-4	Sequence 4, Appl
6	104	100.0	2227	US-09-653-499-2	Sequence 2, Appl
7	104	100.0	2227	US-09-653-499-4	Sequence 4, Appl
8	104	100.0	2227	US-09-653-499-6	Sequence 6, Appl
9	104	100.0	2227	US-09-653-499-8	Sequence 8, Appl
10	104	100.0	2227	US-09-653-499-10	Sequence 10, Appl
11	104	100.0	2227	US-09-653-499-12	Sequence 12, Appl
12	104	100.0	2227	US-09-653-499-14	Sequence 14, Appl
13	104	100.0	2227	US-09-653-499-16	Sequence 16, Appl
14	104	100.0	2227	US-09-653-499-18	Sequence 18, Appl
15	104	100.0	2227	US-09-653-499-20	Sequence 20, Appl
16	104	100.0	2227	US-09-653-499-22	Sequence 22, Appl
17	104	100.0	2227	US-09-653-499-24	Sequence 24, Appl
18	104	100.0	2227	US-09-653-499-26	Sequence 26, Appl
19	104	100.0	2227	US-09-653-499-28	Sequence 28, Appl
20	104	100.0	2227	US-09-653-499-30	Sequence 30, Appl
21	104	100.0	2227	US-09-653-499-32	Sequence 32, Appl
22	104	100.0	2227	US-09-653-499-34	Sequence 34, Appl
23	104	100.0	2227	US-09-653-499-36	Sequence 36, Appl
24	104	100.0	2227	US-09-653-499-38	Sequence 38, Appl
25	104	100.0	2227	US-09-653-499-40	Sequence 40, Appl
26	104	100.0	2227	US-09-653-499-42	Sequence 42, Appl
27	104	100.0	2227	US-09-653-499-44	Sequence 44, Appl

28	39	37.5	609	4	US-09-115-475-28	Sequence 28, Appl
29	39	37.5	609	4	US-09-115-475-31	Sequence 31, Appl
30	39	37.5	1066	4	US-09-004-838-24	Sequence 24, Appl
31	39	37.5	1323	4	US-09-004-838-90	Sequence 90, Appl
32	39	37.5	1604	4	US-09-004-838-95	Sequence 95, Appl
33	39	37.5	1817	4	US-09-004-838-125	Sequence 125, Appl
34	39	37.5	1817	4	US-09-134-001C-3897	Sequence 3897, Ap
35	38	36.5	181	4	US-09-328-352-7035	Sequence 7035, Ap
36	38	36.5	375	4	US-09-252-991A-20086	Sequence 20086, A
37	38	36.5	541	4	US-09-198-452A-692	Sequence 692, App
38	38	36.5	1238	4	US-09-252-991A-26363	Sequence 26363, A
39	38	36.5	1250	3	US-08-938-291A-9	Sequence 9, Appl
40	38	36.5	1250	4	US-08-589-619-9	Sequence 9, Appl
41	38	36.5	1367	2	US-08-249-687C-2	Sequence 2, Appl
42	38	36.5	1367	2	US-08-625-819-2	Sequence 2, Appl
43	38	36.5	1367	2	US-08-746-559A-2	Sequence 2, Appl
44	38	36.5	1367	3	US-08-864-641B-18	Sequence 18, Appl
45	38	36.5	1367	4	US-08-864-641B-18	Sequence 18, Appl

ALIGNMENTS

```

RESULT 1
US-08-475-886-2
Sequence 2, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 08/397,232
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match
Best local Similarity 100.0%; Score 104; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KVNFGMDLEIANSKD 20
DB 922 KVNFGMDLEIANSKD 941
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232

```

```
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 KVNPHGMLDEETIAANSKD 20
Db      922 KVNPHGMLDEETIAANSKD 941
```

```
RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
```

```
GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 KVNPHGMLDEETIAANSKD 20
Db      922 KVNPHGMLDEETIAANSKD 941
```

```
RESULT 4
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
```

```
; TYPE: PRF
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 KVNPHGMLDEETIAANSKD 20
Db      922 KVNPHGMLDEETIAANSKD 941
```

```
RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 KVNPHGMLDEETIAANSKD 20
Db      922 KVNPHGMLDEETIAANSKD 941
```

```
RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; APPLICANT: EMERSON, SUZANNE, U.;
; APPLICANT: PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
```


PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
09-171-387-2

Query Match 100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLLEIANSKD 20
DB 922 KVNPPHGMJDLLEIANSKD 941

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLLEIANSKD 20
DB 922 KVNPPHGMJDLLEIANSKD 941

RESULT 8
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLLEIANSKD 20
DB 922 KVNPPHGMJDLLEIANSKD 941

RESULT 9
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 104; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLLEIANSKD 20
DB 922 KVNPPHGMJDLLEIANSKD 941

RESULT 10
US-09-099-677A-6
Sequence 6, Application US/09099677A
Patent No. 5965369
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lai, Preeti
APPLICANT: Corley, Neil C.

APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN SUCCINYL-COENZYME A SYNTHETASE HOLOENZYME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,677A
FILING DATE: June 18, 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0545 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: g164669
US-09-099-677A-6

Query Match 43.3%; Score 45; DB 2; Length 417;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNPPHMLDEETIANSKD 20
DB 153 VGSFGGVDTIEVVAASNP 171

RESULT 11
US-09-261-471-6
Sequence 6, Application US/09261471
Patent No. 6025123
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/261,471
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/099,677
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0545 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 417 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: g164669
US-09-261-471-6

Query Match 43.3%; Score 45; DB 3; Length 417;
Best Local Similarity 42.1%; Pred. No. 11;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNPPHMLDEETIANSKD 20
DB 153 VGSFGGVDTIEVVAASNP 171

RESULT 12
US-09-099-677A-3
Sequence 3, Application US/09099677A
Patent No. 5965369
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SUCCINYL-COENZYME A SYNTHETASE HOLOENZYME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/099,677A
FILING DATE: June 18, 1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0545 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSBPT06
CLONE: 3273853
US-09-099-677A-3

Query Match 42.3%; Score 44; DB 2; Length 432;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNFPHGMLEETIAANSKD 20
DB 168 VGOPOGVDIEEVAASNP 186

RESULT 13

US-09-261-471-3
Sequence 3, Application US/09261471
Patent No. 6025123
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/261,471
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/099,677
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0545 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSBPT06
CLONE: 3273853
US-09-261-471-3

Query Match 42.3%; Score 44; DB 3; Length 432;
Best Local Similarity 42.1%; Pred. No. 18;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNFPHGMLEETIAANSKD 20
DB 168 VGOPOGVDIEEVAASNP 186

RESULT 14

US-09-181-487-2
Sequence 2, Application US/09181487
Patent No. 6155752
GENERAL INFORMATION:
APPLICANT: DUCKWORTH, DAVID MALCOLM
TITLE OF INVENTION: NOVEL COMPOUNDS

NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/181,487
FILING DATE: 28-OCT-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9807082.4
FILING DATE: 01-APR-1998
APPLICATION NUMBER: 9815489.1
FILING DATE: 16-JUL-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul P.
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30094
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700
TELEX: 846169
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 98 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-181-487-2

Query Match 41.3%; Score 43; DB 3; Length 98;
Best Local Similarity 44.4%; Pred. No. 4.7;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NFPHGMLEETIAANSKD 20
DB 40 NLEPGVADLTQIDVNVQD 57

RESULT 15

US-09-227-357-219
Sequence 219, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
CURRENT FILING DATE: 1999-01-08
EARLIER APPLICATION NUMBER: PCT/US98/13684
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931

Job time : 13.7059 secs

```

: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,932
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,916
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,930
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,918
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,920
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/052,733
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/052,795
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,919
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/051,928
: EARLIER FILING DATE: 1997-07-08
: EARLIER APPLICATION NUMBER: 60/055,722
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,723
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,948
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,949
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,953
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,950
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,947
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,964
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/056,360
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,684
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,984
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/055,954
: EARLIER FILING DATE: 1997-08-18
: EARLIER APPLICATION NUMBER: 60/058,785
: EARLIER FILING DATE: 1997-09-12
: EARLIER APPLICATION NUMBER: 60/058,664
: EARLIER FILING DATE: 1997-09-12
: EARLIER APPLICATION NUMBER: 60/058,660
: EARLIER FILING DATE: 1997-09-12
: EARLIER APPLICATION NUMBER: 60/058,661
: EARLIER FILING DATE: 1997-09-12
: NUMBER OF SEQ ID NOS: 672
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 219
: LENGTH: 99
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (99)
: OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-219

```

```

Query Match      41.3%; Score 43; DB 4; Length 99;
Best Local Similarity 44.4%; Pred. No. 4.7;
Matches      8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

OY      3 NPPHGMLEIEIANSKD 20
      1 | | | | | | | | | |
Db      40 NLEPGVADLTQIDVNVOD 57

```

Search completed: October 1, 2003, 10:06:31

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 : Search time 24.1176 Seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-44
Perfect score: 104
Sequence: 1 KVNFPFGMDLEETIANSKD 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	2227	10	US-09-929-955-12
2	104	100.0	2227	14	US-10-104-966-12
3	104	100.0	2227	14	US-10-135-988-2
4	104	100.0	2227	14	US-10-135-988-4
5	104	100.0	2227	14	US-10-135-988-6
6	98	94.2	352	15	US-10-272-459-45
7	98	94.2	980	15	US-10-272-459-41
8	43	41.3	98	10	US-09-978-255A-616
9	43	41.3	98	10	US-09-978-657-616
10	43	41.3	98	10	US-09-978-192A-616
11	43	41.3	98	11	US-09-999-832A-616
12	43	41.3	98	11	US-09-978-189-616
13	43	41.3	98	11	US-09-978-608A-616
14	43	41.3	98	11	US-09-978-585A-616
15	43	41.3	98	11	US-09-978-191A-616

16	43	41.3	98	11	US-09-978-403A-616	Sequence 616, App
17	43	41.3	98	11	US-09-978-564A-616	Sequence 616, App
18	43	41.3	98	11	US-09-999-833A-616	Sequence 616, App
19	43	41.3	98	11	US-09-981-915A-616	Sequence 616, App
20	43	41.3	98	11	US-09-978-824-616	Sequence 616, App
21	43	41.3	98	11	US-09-918-585A-616	Sequence 616, App
22	43	41.3	98	11	US-09-978-423A-616	Sequence 616, App
23	43	41.3	98	11	US-09-978-193A-616	Sequence 616, App
24	43	41.3	98	11	US-09-999-830A-616	Sequence 616, App
25	43	41.3	98	11	US-09-978-757A-616	Sequence 616, App
26	43	41.3	98	11	US-09-978-187B-616	Sequence 616, App
27	43	41.3	98	11	US-09-978-643A-616	Sequence 616, App
28	43	41.3	98	12	US-09-978-375A-616	Sequence 616, App
29	43	41.3	98	12	US-09-978-188A-616	Sequence 616, App
30	43	41.3	98	12	US-09-978-298A-616	Sequence 616, App
31	43	41.3	98	12	US-10-143-031A-616	Sequence 616, App
32	43	41.3	98	12	US-10-002-967A-616	Sequence 616, App
33	43	41.3	98	12	US-10-017-083A-616	Sequence 616, App
34	43	41.3	98	12	US-10-143-030A-616	Sequence 616, App
35	43	41.3	98	12	US-10-216-163-106	Sequence 106, App
36	43	41.3	98	12	US-10-145-128A-616	Sequence 616, App
37	43	41.3	98	12	US-10-017-191A-616	Sequence 616, App
38	43	41.3	98	12	US-10-143-028A-616	Sequence 616, App
39	43	41.3	98	12	US-10-143-029A-616	Sequence 616, App
40	43	41.3	98	12	US-10-145-089A-616	Sequence 616, App
41	43	41.3	98	15	US-10-227-884-106	Sequence 106, App
42	43	41.3	98	15	US-10-230-163-106	Sequence 106, App
43	43	41.3	98	15	US-10-230-338-106	Sequence 106, App
44	43	41.3	98	15	US-10-218-631-106	Sequence 106, App
45	43	41.3	98	15	US-10-017-081A-616	Sequence 616, App

ALIGNMENTS

RESULT 1
US-09-929-955-12
: Sequence 12, Application US/09929955
: Patent No. US20020136740A1
: GENERAL INFORMATION:
: APPLICANT: Matti Sallberg
: TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
: FILE REFERENCE: TRIPEP.23AUS2
: CURRENT APPLICATION NUMBER: US/09/929,955
: PRIOR FILING DATE: 2001-08-15
: PRIOR APPLICATION NUMBER: 09/705,547
: PRIOR FILING DATE: 2000-11-03
: PRIOR APPLICATION NUMBER: 60/229,175
: PRIOR FILING DATE: 2000-08-29
: PRIOR APPLICATION NUMBER: 60/225,767
: PRIOR FILING DATE: 2000-08-17
: NUMBER OF SEQ ID NOS: 49
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12
: LENGTH: 2227
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match Best Local Similarity 100.0%: Score 104; DB 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMDLEETIANSKD 20
DB 922 KVNFPFGMDLEETIANSKD 941

RESULT 2

US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT FILING DATE: US/10/104,966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 104; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETANSKD 20
|||||
Db 922 KVNPHGMLDEETANSKD 941

RESULT 3
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT FILING DATE: US/10/135,988
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 104; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETANSKD 20
|||||
Db 922 KVNPHGMLDEETANSKD 941

RESULT 4
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT FILING DATE: US/10/135,988
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 104; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETANSKD 20
|||||
Db 922 KVNPHGMLDEETANSKD 941

RESULT 5
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT FILING DATE: US/10/135,988
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 104; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1,7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETANSKD 20
|||||
Db 922 KVNPHGMLDEETANSKD 941

RESULT 6
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: Ppl7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459

CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match 94.2%; Score 98; DB 15; Length 352;
Best Local Similarity 95.0%; Pred. No. 2.1e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KVNPHGMJLDEIANSKD 20
||||| |||||||||
294 KVNPHAMJLDEIANSKD 313

US-10-272-459-41
Sequence 41, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANTES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: P017955 002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 41
LENGTH: 980
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match 94.2%; Score 98; DB 15; Length 980;
Best Local Similarity 95.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KVNPHGMJLDEIANSKD 20
||||| |||||||||
922 KVNPHAMJLDEIANSKD 941

RESULT 8
US-09-978-295A-616
Sequence 616, Application US/09978295A
Patent No. US2002015606A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fond, Sherman
APPLICANT: Geo, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.

APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James:
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/07450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327

;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06

;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NFPHGMDLEIRANSKD 20
DB 40 NLEPGVADLTQIDVNOVD 57

RESULT 9
US-978-697-616
Sequence 616, Application US/09978697
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083316
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598

PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-5-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPPHGMDLERIAANSKD 20
Db 40 NLEPGVADLQIDVNVOD 57

RESULT 10
US-09-978-192A-616
Sequence 616, Application US/09978192A
Patent No. US2002017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C9

CURRENT APPLICATION NUMBER: US/09/978,192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195

PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPPHGMLEEEIRANSKD 20
DB 40 NIPEGVADLTQIDVNVQD 57

RESULT 11

US-09-999-832a-616

Sequence 616, Application US/09999832A

Publication No. US20020192706A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlisen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Klavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paonli, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC63
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249

1 PRIOR FILING DATE: 1997-11-03
 2 PRIOR APPLICATION NUMBER: 60/065311
 3 PRIOR FILING DATE: 1997-11-13
 4 PRIOR APPLICATION NUMBER: 60/063644
 5 PRIOR FILING DATE: 1997-11-21
 6 PRIOR APPLICATION NUMBER: 60/077450
 7 PRIOR FILING DATE: 1996-03-10
 8 PRIOR APPLICATION NUMBER: 60/076532
 9 PRIOR FILING DATE: 1996-03-11
 10 PRIOR APPLICATION NUMBER: 60/076411
 11 PRIOR FILING DATE: 1996-03-11
 12 PRIOR APPLICATION NUMBER: 60/077649
 13 PRIOR FILING DATE: 1996-03-11
 14 PRIOR APPLICATION NUMBER: 60/077791
 15 PRIOR FILING DATE: 1996-03-12
 16 PRIOR APPLICATION NUMBER: 60/078004
 17 PRIOR FILING DATE: 1996-03-13
 18 PRIOR APPLICATION NUMBER: 60/078886
 19 PRIOR FILING DATE: 1996-03-20
 20 PRIOR APPLICATION NUMBER: 60/078936
 21 PRIOR FILING DATE: 1996-03-20
 22 PRIOR APPLICATION NUMBER: 60/078910
 23 PRIOR FILING DATE: 1996-03-20
 24 PRIOR APPLICATION NUMBER: 60/078939
 25 PRIOR FILING DATE: 1996-03-20
 26 PRIOR APPLICATION NUMBER: 60/079294
 27 PRIOR FILING DATE: 1996-03-25
 28 PRIOR APPLICATION NUMBER: 60/079656
 29 PRIOR FILING DATE: 1996-03-26
 30 PRIOR APPLICATION NUMBER: 60/079664
 31 PRIOR FILING DATE: 1996-03-27
 32 PRIOR APPLICATION NUMBER: 60/079689
 33 PRIOR FILING DATE: 1996-03-27
 34 PRIOR APPLICATION NUMBER: 60/079663
 35 PRIOR FILING DATE: 1996-03-27
 36 PRIOR APPLICATION NUMBER: 60/079728
 37 PRIOR FILING DATE: 1996-03-27
 38 PRIOR APPLICATION NUMBER: 60/079786
 39 PRIOR FILING DATE: 1996-03-27
 40 PRIOR APPLICATION NUMBER: 60/079920
 41 PRIOR FILING DATE: 1996-03-30
 42 PRIOR APPLICATION NUMBER: 60/079923
 43 PRIOR FILING DATE: 1996-03-30
 44 PRIOR APPLICATION NUMBER: 60/080105
 45 PRIOR FILING DATE: 1996-03-31
 46 PRIOR APPLICATION NUMBER: 60/080107
 47 PRIOR FILING DATE: 1996-03-31
 48 PRIOR APPLICATION NUMBER: 60/080165
 49 PRIOR FILING DATE: 1996-03-31
 50 PRIOR APPLICATION NUMBER: 60/080194
 51 PRIOR FILING DATE: 1996-03-31
 52 PRIOR APPLICATION NUMBER: 60/080327
 53 PRIOR FILING DATE: 1996-04-01
 54 PRIOR APPLICATION NUMBER: 60/080328
 55 PRIOR FILING DATE: 1996-04-01
 56 PRIOR APPLICATION NUMBER: 60/080333
 57 PRIOR FILING DATE: 1996-04-01
 58 PRIOR APPLICATION NUMBER: 60/080334
 59 PRIOR FILING DATE: 1996-04-01
 60 PRIOR APPLICATION NUMBER: 60/081070
 61 PRIOR FILING DATE: 1996-04-08
 62 PRIOR APPLICATION NUMBER: 60/081049
 63 PRIOR FILING DATE: 1996-04-08
 64 PRIOR APPLICATION NUMBER: 60/081071
 65 PRIOR FILING DATE: 1996-04-08
 66 PRIOR APPLICATION NUMBER: 60/081195
 67 PRIOR FILING DATE: 1996-04-08
 68 PRIOR APPLICATION NUMBER: 60/081203
 69 PRIOR FILING DATE: 1996-04-09
 70 PRIOR APPLICATION NUMBER: 60/081229
 71 PRIOR FILING DATE: 1996-04-09
 72 PRIOR APPLICATION NUMBER: 60/081955
 73 PRIOR FILING DATE: 1996-04-15

1	PRIOR APPLICATION NUMBER: 60/081817
2	PRIOR FILING DATE: 1998-04-15
3	PRIOR APPLICATION NUMBER: 60/081819
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081952
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/081838
8	PRIOR FILING DATE: 1998-04-15
9	PRIOR APPLICATION NUMBER: 60/082568
10	PRIOR FILING DATE: 1998-04-21
11	PRIOR APPLICATION NUMBER: 60/082569
12	PRIOR FILING DATE: 1998-04-21
13	PRIOR APPLICATION NUMBER: 60/082704
14	PRIOR FILING DATE: 1998-04-22
15	PRIOR APPLICATION NUMBER: 60/082804
16	PRIOR FILING DATE: 1998-04-22
17	PRIOR APPLICATION NUMBER: 60/082700
18	PRIOR FILING DATE: 1998-04-22
19	PRIOR APPLICATION NUMBER: 60/082797
20	PRIOR FILING DATE: 1998-04-22
21	PRIOR APPLICATION NUMBER: 60/082796
22	PRIOR FILING DATE: 1998-04-23
23	PRIOR APPLICATION NUMBER: 60/083336
24	PRIOR FILING DATE: 1998-04-27
25	PRIOR APPLICATION NUMBER: 60/083322
26	PRIOR FILING DATE: 1998-04-28
27	PRIOR APPLICATION NUMBER: 60/083392
28	PRIOR FILING DATE: 1998-04-29
29	PRIOR APPLICATION NUMBER: 60/083495
30	PRIOR FILING DATE: 1998-04-29
31	PRIOR APPLICATION NUMBER: 60/083496
32	PRIOR FILING DATE: 1998-04-29
33	PRIOR APPLICATION NUMBER: 60/083499
34	PRIOR FILING DATE: 1998-04-29
35	PRIOR APPLICATION NUMBER: 60/083545
36	PRIOR FILING DATE: 1998-04-29
37	PRIOR APPLICATION NUMBER: 60/083544
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/083558
40	PRIOR FILING DATE: 1998-04-29
41	PRIOR APPLICATION NUMBER: 60/083539
42	PRIOR FILING DATE: 1998-04-29
43	PRIOR APPLICATION NUMBER: 60/083500
44	PRIOR FILING DATE: 1998-04-29
45	PRIOR APPLICATION NUMBER: 60/083742
46	PRIOR FILING DATE: 1998-04-30
47	PRIOR APPLICATION NUMBER: 60/084366
48	PRIOR FILING DATE: 1998-05-05
49	PRIOR APPLICATION NUMBER: 60/084414
50	PRIOR FILING DATE: 1998-05-06
51	PRIOR APPLICATION NUMBER: 60/084441
52	PRIOR FILING DATE: 1998-05-06
53	PRIOR APPLICATION NUMBER: 60/084637
54	PRIOR FILING DATE: 1998-05-07
55	PRIOR APPLICATION NUMBER: 60/084639
56	PRIOR FILING DATE: 1998-05-07
57	PRIOR APPLICATION NUMBER: 60/084640
58	PRIOR FILING DATE: 1998-05-07
59	PRIOR APPLICATION NUMBER: 60/084598
60	PRIOR FILING DATE: 1998-05-07
61	PRIOR APPLICATION NUMBER: 60/084600
62	PRIOR FILING DATE: 1998-05-07
63	PRIOR APPLICATION NUMBER: 60/084627
64	PRIOR FILING DATE: 1998-05-07
65	PRIOR APPLICATION NUMBER: 60/084643
66	PRIOR FILING DATE: 1998-05-07
67	PRIOR APPLICATION NUMBER: 60/085339
68	PRIOR FILING DATE: 1998-05-13
69	PRIOR APPLICATION NUMBER: 60/085338
70	PRIOR FILING DATE: 1998-05-13
71	PRIOR APPLICATION NUMBER: 60/085323
72	PRIOR FILING DATE: 1998-05-13
73	PRIOR APPLICATION NUMBER: 60/085582

[illegible]

? PRIOR APPLICATION NUMBER: 60/085560
 ? PRIOR FILING DATE: 1998-05-15
 ? PRIOR APPLICATION NUMBER: 60/085573
 ? PRIOR FILING DATE: 1998-05-15
 ? PRIOR APPLICATION NUMBER: 60/085704
 ? PRIOR FILING DATE: 1998-05-15
 ? PRIOR APPLICATION NUMBER: 60/085697

Query Match	41.3%	Score 43;	DB 11;	Length 98;
Best Local Similarity	44.4%;	Pred. NO 9.9;		
Matches	8;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      3 NEPHGMLDLEEIAANSKD 20
        | | | : | | : |
Db      40 NLEPGVADLTQIDVNVQD 57
```

RESULT 13
HE-00-078

```

US-09-978-608A-616
: Sequence 616, Application US/09978608A
: Publication No. US20030045462A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Botstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleon
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gerritsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Hillan, Kenneth J.
: APPLICANT: Kljavin, Ivar J.
: APPLICANT: Kuo, Sophia S.
: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James?
: APPLICANT: Paonli, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630P1C22
: CURRENT APPLICATION NUMBER: US/09/978,608A
: NUMBER OF SEQ ID NOS: 624
: Prior Application removed - See File Wrapper or Palm
: SEQ ID NO 616
: LENGTH: 98
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-978-608A-616

```

Query Match	41.3%	Score 43:	DB 11;	Length 98;
Best Local Similarity	44.4%;	Pred. No. 9.9;		
Matches 8;	Conservative 3;	Mismatches 7;	Indels 0;	Gaps 0;

```
QY      3 NFRHGMLDLEEIAANSKD 20
        | | | : | | | : |
Db      40 NLREGVADLTQIDVNVQD 57
```

RESULT 14
US-09-978-585A-616

```

Sequence 616, Application US/09978585A
Publication No. US20030049633A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Deanoysers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Olang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C15
CURRENT APPLICATION NUMBER: US/09/978,585A
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ ID NOS: 624
Prior Application removed - See File Wrapper or Palm
SEQ ID NO 616
LENGTH: 98
TYPE: PRT
ORGANISM: Homo Sapien
US-09-978-585A-616
Query Match 41.3%; Score 43; DB 11; Length 98;
Best Local Similarity 44.4%; Pred. No. 9.9;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
3 NEPHGMIDLEETIANSKD 20
| | | | | | | | | |
40 NLPEGVADLTQIDVNVOD 57
RESULT 5
US-09-978-191A-616
Sequence 616, Application US/09978191A
Publication No. US20030050239A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Deanoysers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Olang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.

```

APPLICANT: Kijavni, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC4
CURRENT APPLICATION NUMBER: US/09/978,191A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918555
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077652
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077921
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079669
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079778
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080337

PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080328
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080333
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/080334
 PRIOR FILING DATE: 1998-04-01
 PRIOR APPLICATION NUMBER: 60/081070
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081049
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081071
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081195
 PRIOR FILING DATE: 1998-04-08
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081955
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081819
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081952
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081838
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082568
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082569
 PRIOR FILING DATE: 1998-04-21
 PRIOR APPLICATION NUMBER: 60/082704
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082804
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082700
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082797
 PRIOR FILING DATE: 1998-04-22
 PRIOR APPLICATION NUMBER: 60/082796
 PRIOR FILING DATE: 1998-04-23
 PRIOR APPLICATION NUMBER: 60/083336
 PRIOR FILING DATE: 1998-04-27
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083392
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083495
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083496
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083499
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083554
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083558
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083559
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083500
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/083742
 PRIOR FILING DATE: 1998-04-30
 PRIOR APPLICATION NUMBER: 60/084366
 PRIOR FILING DATE: 1998-05-05
 PRIOR APPLICATION NUMBER: 60/084414
 PRIOR FILING DATE: 1998-05-06
 PRIOR APPLICATION NUMBER: 60/084441
 PRIOR FILING DATE: 1998-05-06

PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084639
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084640
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084598
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084643
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085582
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085700
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085689
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085580
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085573
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.38; Score 43; DB 11; Length 98;
 Best Local Similarly 44.48; Pred. No. 9.9;
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 Oy 3 NPPGMDLEFIANSKD 20
 Db 40 NPEGVADLTQIDVYVD 57

Search completed: October 1, 2003, 10:37:52
 Job time : 25.1176 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-45
Perfect score: 101
Sequence: 1 DDEETANSKDPNNSEFDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	101	100.0	20	AAW42928	Immunogenic Hepati
2	101	100.0	21	ABG31729	Synthetic HAV P2A
3	101	100.0	22	ABG31729	Sequence of viral
4	101	100.0	22	ABG31729	Attenuated Hepatit
5	101	100.0	22	ABG31729	Hepatitis A virus
6	101	100.0	22	ABG31729	Amino acid sequenc
7	101	100.0	22	ABG31729	Amino acid sequenc
8	101	100.0	22	ABG31729	Wild-type Hepatiti
9	101	100.0	22	ABG31727	

10	101	100.0	22	ABG31728	Hepatitis A virus
11	101	100.0	22	ABG31729	Attenuated Hepatit
12	101	100.0	22	ABG31729	Hepatitis A virus
13	101	100.0	22	ABG31729	Wild type human he
14	101	100.0	22	ABG31729	Attenuated (pass3
15	101	100.0	22	ABG31729	Immunogenic Hepati
16	101	100.0	22	ABG31729	Synthetic HAV P2A
17	53	52.5	20	AAW42927	Drosophila melanog
18	46	45.5	21	ABG31729	Drosophila melanog
19	44.5	44.1	17	ABG31729	Drosophila melanog
20	44	43.6	80	ABG31729	Drosophila melanog
21	43	42.6	152	ABG31729	Human ORF protein
22	43	42.6	157	ABG31729	H. pylori ORF hp6
23	43	42.6	350	ABG31729	H. pylori ORF hp2P
24	43	42.6	20	AAV27183	(S)-3'-hydroxy-N-m
25	43	42.6	536	ABG31729	Arabidopsis thalia
26	43	42.6	1201	AAW90345	Drosophila sp. Cos
27	43	42.6	1201	ABG31729	Drosophila melanog
28	42	41.6	89	AAV5799	Propionibacterium
29	42	41.6	177	AAV48507	Arabidopsis thalia
30	42	41.6	226	AAV48505	Arabidopsis thalia
31	42	41.6	265	AAV48505	Arabidopsis thalia
32	42	41.6	365	AAV29461	Pentactinaria thali
33	42	41.6	402	ABG31729	Enterobacter cloac
34	42	41.6	425	ABG31729	Drosophila melanog
35	42	41.6	439	ABP78252	Thermus caldophilu
36	42	41.6	445	AAW40763	N. gonorrhoeae aml
37	42	41.6	445	AAW40764	Human polypeptide
38	42	41.6	445	AAW40765	Human polypeptide
39	42	41.6	538	AAV81713	Streptococcus pneu
40	42	41.6	538	ABU00469	S. pneumoniae type
41	42	41.6	641	AAW38978	Human polypeptide
42	42	41.6	679	AAO17073	Human t2 protein
43	42	41.6	698	AAW38977	Human polypeptide
44	42	41.6	698	AAW38977	Human polypeptide
45	42	41.6	705	AAW94135	Human protein sequ

ALIGNMENTS

RESULT 1
ID AAW42928 standard; peptide: 20 AA.
XX AAW42928;
AC
AC AAW42928;
XX 28-APR-1998 (first entry)
DT
XX Immunogenic Hepatitis A virus peptide YK-1328.
DE
XX Immunogenic Hepatitis A virus.
XX Immune response; antibody.
XX Synthetic.
OS Hepatitis A virus.
OS
PN WO9740147-A1.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06891.
XX
XX 19-APR-1996; 96US-0015644.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE;
XX Fields HA, Khudyakov YE;
XX WPI: 1997-535831/49.
XX
XX Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an
XX Immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAM42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 931-950. The present peptide
CC is derived from amino acids 931-950, and has a reactivity of 12.5% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC administering the peptide to a mammal.
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 101; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DLEIRIANSKDFPMSETDL 20
1 DLEIRIANSKDFPMSETDL 20
Db
RESULT 2
AAB69445
ID AAB69445 standard; Peptide: 21 AA.
XX
AC AAB69445;
XX
DT 20-APR-2001 (first entry)
DE Synthetic HAV P2A peptide, SEQ ID NO: 45.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
PF 15-JUL-1999; 99US-0144412.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines -
PS Claim 13; Page 97; 130pp; English.
XX
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
SQ Sequence 21 AA;
Query Match 100.0%; Score 101; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DLEIRIANSKDFPMSETDL 20
1 DLEIRIANSKDFPMSETDL 20
Db
RESULT 3
AAP60066
ID AAP60066 standard; Protein; 2227 AA.
XX
AC AAP60066;
XX
DT 25-MAR-2003 (updated)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of viral I434 polypeptide encoded by the complete
XX nucleotide sequence of the HAV genome.
XX
KW Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
XX
XX Key Location/Qualifiers
XX Region 1..245
XX /label= P1.1A
XX Region 246..491
XX /label= 1B
XX Region 492..836
XX /label= 1C
XX Region 837..980
XX /label= P2.2A
XX Region 981..1076
XX /label= 2B
XX Region 1077..1422
XX /label= 2C
XX Region 1423..1484
XX /label= P3.3A
XX Region 1485..1507
XX /label= 3B
XX Region 1508..1678
XX /label= 3C
XX Region 1679..2227
XX /label= 3D
XX
XX EPI99480-A.
XX
XX 29-OCT-1986.
XX
XX 03-APR-1986; 86EP-0302465.
XX
XX 03-APR-1985; 85US-0719329.
XX
XX (CHIR) CHIRON CORP.
XX
XX Dina D, Potter SJ, Vannest GA, Caput D;
XX WPI; 1986-286213/44.
XX N-PSDB; AAN60080.
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use
XX in prodn. of vaccines and diagnostic probes
XX
PS Claim 5; Fig 1; 18pp; English.

XX	AAN6080 and oligonucleotide fragments are useful in detection of
CC	hepatitis A virus; transformed hosts may be used for expression of
CC	polypeptides and fragments useful in vaccines without risk of
CC	infection by the virus or in prodn. of particles which are capable
CC	of inducing immunocompetent B cells for passive immunotherapy. Pref.
CC	epitope is derived from AAs 445-657 or 792-848 of the HAV
CC	polypeptide sequence (AAP60066).
CC	(Updated on 25-MAR-2003 to correct PA field.)
SO	Sequence 2227 AA;
OY	Query Match 100.0%; Score 101; DB 7; Length 2227; Best Local Similarity 100.0%; Pred. NO. 2.5e-07; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
DB	1 DLEEIAANSKDFPNMSETDL 20 931 DLEEIAANSKDFPNMSETDL 950
SOURCE	AAAR05697 standard; protein; 2227 AA. AAR05697
AC	AAR05697;
DT	25-MAR-2003 (updated)
DT	15-AUG-1990 (first entry)
XX	Attenuated hepatitis A virus.
DE	Hepatitis A virus; vaccine; attenuated.
KW	Hepatitis A virus; strain HM-175.
OS	Hepatitis A virus, strain HM-175.
XX	
FH	Key Location/Qualifiers
FT	Region 1..23
FT	/label=VP4 - 1A
FT	24..245
FT	/label=VP2 - 1B
FT	246..491
FT	/label=VP3 - 1C
FT	492..791
FT	/label=VP1 - 1D
FT	792..980
FT	/label=2A
FT	981..1087
FT	/label=2B
FT	1088..1422
FT	/label=2C
FT	1423..1496
FT	/label=3A
FT	1497..1519
FT	/label=3B - VPg
FT	1520..1738
FT	/label=3C
FT	1739..2227
FT	/label=3D
XX	
PN	US4894228-A.
XX	
PD	16-JAN-1990.
XX	
PE	12-JUL-1988; 88US-0217824.
XX	
PR	12-JUL-1988; 88US-0217824.
XX	
PA	12-JUL-1988; 88US-0652967.
XX	
TI	(USSH) US DEPT HEALTH & HUMAN SERVICE.
TI	Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM, Daemer KU, Gust ID;

```

xx WP1: 1990-075557/10.
DR N-PSDB: AA003512.
xx
xx Vaccine against hepatitis A virus infection - comprises novel
xx attenuated hepatitis A virus strain.
xx
xx Claim 1; Fig 1; 18pp; English.
xx
xx The attenuated HAV is useful for inducing protective immunity against
xx HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
xx several nucleotide changes distributed throughout the genome, is
xx attenuated for chimpanzees, elicits serum neutralising antibodies, and is
xx suitable for use as an HAV vaccine. It is noted that not all the changes
xx are necessary for attenuation and use as a vaccine.
xx (Updated on 25-MAR-2003 to correct PA field.)
xx
xx Sequence 2227 AA;
xx
xx Query Match 100.0%; Score 101; DB 11; Length 2227;
xx Best Local Similarity 100.0%; Pred. No. 2.5e-07;
xx Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
xx
xx 1 DLEERAAASKDPNNSETDL 20
xx ||||||||||||||||
xx
xx Db 931 DLEERAAASKDPNNSETDL 950
xx
xx RESULT 5
xx AAWM34074
xx AAWM34074 standard; Protein; 2227 AA.
xx
xx AAWM34074;
xx
xx 27-APR-1998 (first entry)
xx
xx Hepatitis A virus HM-175 protein sequence.
xx
xx HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
xx infection; vaccine.
xx
xx Hepatitis A virus HM-175.
xx
xx Key Location/Qualifiers
xx FH 1..23
xx FH Protein /label= VP4
xx FT 24..245
xx FT /label= VP2
xx FT 246..491
xx FT /label= VP3
xx FT 492..791
xx FT /label= VP1
xx FT 792..980
xx FT /label= 2A
xx FT 981..1087
xx FT /label= 2B
xx FT 1088..1422
xx FT /label= 2C
xx FT 1423..1496
xx FT /label= 3A
xx FT 1497..1519
xx FT /label= 3B
xx FT 1520..1738
xx FT /label= 3C
xx FT 1739..2227
xx FT /label= 3D
xx
xx MO9740166-A2.
xx
xx 30-OCT-1997.
xx
xx 18-APR-1997; 97WO-US06506.
xx

```

XX 19-APR-1996; 960S-0015642.
XX (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Emerson SU, Purcell RH, Raychaudhuri G;
PI WPI; 1997-535850/49.
DR N-PSDB; AAT93023.
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
PT as vaccines against HAV infection
XX Disclosure: Fig 13A-D; 66pp; English.
XX This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (1) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DLEBIAANSKDFPMSETDL 20
DB 931 DLEBIAANSKDFPMSETDL 950
RESULT 6
AAB18607
ID AAB18607 standard; Protein: 2227 AA.
XX AAB18607;
XX 15-JAN-2001 (first entry)
XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX Hepatitis A virus.
XX OS US6113912-A.
XX PN US6113912-A.
XX PD 05-SEP-2000.
XX PF 07-JUN-1995; 950S-0475886.
XX PR 18-SEP-1992; 920S-0947338.
XX PR 17-SEP-1993; 93MO-US08610.
XX PR 10-MAR-1995; 950S-0397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.

DR N-PSDB; AAA75476.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX XX
XX Disclosure: Fig 6A-K; 72pp; English.
XX The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DLEBIAANSKDFPMSETDL 20
DB 931 DLEBIAANSKDFPMSETDL 950
RESULT 7
AAB18608
ID AAB18608 standard; Protein: 2227 AA.
XX AAB18608;
XX 15-JAN-2001 (first entry)
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX P-35 virus.
XX OS Hepatitis A virus.
XX PN US6113912-A.
XX PD 05-SEP-2000.
XX PF 07-JUN-1995; 950S-0475886.
XX PR 18-SEP-1992; 920S-0947338.
XX PR 17-SEP-1993; 93MO-US08610.
XX PR 10-MAR-1995; 950S-0397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75477.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX XX
XX Disclosure: Columns 67-78; 72pp; English.
XX The present sequence is derived from passage 35 of a wild type
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
CC designated P-35 virus. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
 DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 8
 AAB18609
 ID AAB18609 standard; Protein; 2227 AA.

AC AAB18609;
 DT 15-JAN-2001 (first entry)
 DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
 HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 HAV 4380.

OS Hepatitis A virus.
 PN US6113912-A.
 PD 05-SEP-2000.
 PF 07-JUN-1995; 95US-0475886.
 PR 18-SEP-1992; 92US-0947338.
 PR 17-SEP-1993; 93MO-US08610.
 PR 10-MAR-1995; 95US-0397232.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 DR WPI; 2000-586464/55.
 DR N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type
 PS Disclosure; Columns 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A
 CC virus (HAV) of the invention, designated HAV 4380. The sequence is
 CC produced by modifying wild type HAV strain HM-174. The HAV of the
 CC invention are adapted to growth in the human fibroblast-like cell
 CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
 CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
 CC of hepatitis A in humans and other primates.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
 DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 9
 ABG31727
 ID ABG31727 standard; Protein; 2227 AA.
 AC ABG31727;
 FT

DT 29-NOV-2002 (first entry)
 DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
 XX
 XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
 OS Hepatitis A virus strain HM-175.
 OS US6423318-B1.

PN 23-JUL-2002.

PF 31-AUG-2000; 2000US-0653499.
 PR 07-JUN-1995; 95US-0475886.
 PR 17-SEP-1993; 93US-0397232.
 PR 17-SEP-1993; 93MO-US08610.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
 DR WPI; 2002-680946/73.
 DR N-PSDB; ABS52787.

PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection
 PS Disclosure; Fig 6; 71pp; English.

XX The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC strain HM-175 polypeptide.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
 DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 10
 ABG31728
 ID ABG31728 standard; Protein; 2227 AA.
 AC ABG31728;
 FT

DT 29-NOV-2002 (first entry)

DE Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.

XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
 KW virucide; mutant; pHAV/7; mutain.

OS Hepatitis A virus strain HM-175.
 OS Synthetic.

PH Key Location/Qualifiers

FT Misc-difference 963 /Label= Wild-type Lys substituted by Arg

FT Misc-difference 764 /note= Wild-type Glu substituted by Val

FT Misc-difference 821 /note= Wild-type Asn substituted by Ser

FT Misc-difference 1052 /note= Wild-type Ala substituted by Val

```
FT Misc-difference /note= "Wild-type Gly substituted by Ala"  
ET FT Misc-difference 1062  
ET FT Misc-difference 1118  
ET FT Misc-difference /note= "Wild-type Lys substituted by Met"  
ET FT Misc-difference 1151  
ET FT Misc-difference /note= "Wild-type Glu substituted by Lys"  
ET FT Misc-difference 1163  
ET FT Misc-difference /note= "Wild-type Phe substituted by Ser"  
ET FT Misc-difference 1277  
ET FT Misc-difference /note= "Wild-type Val substituted by Ile"  
ET FT Misc-difference 1500  
ET FT Misc-difference /note= "Wild-type His substituted by Tyr"  
ET FT Misc-difference 1805  
ET FT Misc-difference /note= "Wild-type Asp substituted by Asn"  
ET FT Misc-difference 1930  
ET FT Misc-difference /note= "Wild-type Ser substituted by Thr"  
XX US6423318-B1.  
XX XX  
XX 23-JUL-2002.  
XX 31-AUG-2000; 2000US-0653499.  
XX 07-JUN-1995; 95US-0475886.  
PR 17-SEP-1993; 93US-0397232.  
PR 17-SEP-1993; 93WO-US08610.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
XX WPI: 2002-680946/73.  
DR N-PSDB; ABS52788.  
XX  
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -  
XX  
XX Example 3: Column 67-78; 71pp; English.  
PS The invention relates to a polynucleotide which encodes a hepatitis A  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
CC line). The polynucleotide is useful for preparing a vaccine against  
CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
CC mutant strain HM-175/7 (PHAV/7) polypeptide.  
CC  
SQ Sequence 2227 AA:  
  
Very Match 100.0%; Score 101; DB 23; Length 2227;  
Last Local Similarity 100.0%; Pred. No. 2.Se-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DLERIANSKDPNNSETDL 20  
db |||||  
931 DLERIANSKDFPNNSETDL 950  
  
RESULT 11  
ABG31729  
ID ABG31729 standard; Protein; 2227 AA.  
XX  
XX ABG31729;  
XX  
XX 29-NOV-2002 (first entry)  
DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.  
XX  
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;  
KW HAV 4380.  
XX  
XX Hepatitis A virus strain HM-175.  
OS  
EN US6423318-B1
```

```

XX      23-JUL-2002.
PD      31-AUG-2000; 2000US-0653499.
XX      XX
PF      07-JUN-1995;    95US-0475886.
PR      17-SEP-1993;    93US-0397232.
XX      PR      17-SEP-1993;    93WO-US08610.
PA      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA      (SMIT ) SMITHKLINE BEECHAM BIOLOGICALS.
XX      PI      Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX      DR      WPI: 2002-680946/73.
XX      DR      N-PSDB: ABS52789.
XX      PT      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX      PS      In MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX      CC      Disclosure; Column 93-104; 71pp; English.
XX      CC      The invention relates to a polynucleotide which encodes a hepatitis A
XX      CC      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX      CC      line). The polynucleotide is useful for preparing a vaccine against
XX      CC      hepatitis A virus infection. This sequence represents an attenuated
XX      CC      hepatitis A virus 4830 polypeptide.
SO      Sequence      2227 AA;

Query Match          100.0%; Score 101; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 DLEETIANSKDPPNMSETDL 20
        |||||||
Db       931 DLEETIANSKDPPNMSETDL 950

RESULT 12
AAEI9899
ID      AAEI9899 standard; Protein; 2227 AA.
XX      AC      AAEI9899;
XX      DT      18-JUN-2002 (first entry)
XX      DE      Hepatitis A virus (HAV) protein.
XX      KW      Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX      KW      cytoskeletal; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX      OS      Hepatitis A virus.
XX      PN      WO200213855-A2.
XX      PD      21-FEB-2002.
XX      PF      15-AUG-2001; 2001WO-IB01808.
XX      PR      17-AUG-2000; 2000US-225767P.
XX      PR      29-AUG-2000; 2000US-229175P.
XX      PR      03-NOV-2000; 2000US-0705547.
XX      PA      (TRIP-) TRIPEP AB.
XX      PI      Saliberg M, Hultgren C;
XX      DR      WPI: 2002-241837/29.
XX      DR      N-PSDB: AAD31766.
XX      PT      Vaccine compositions for treating and preventing disease, preferably
XX      PT      hepatitis C virus infection, comprises ribavirin and antigen that has

```

PT epitope present in hepatitis C virus -

XX Claim 11; Page 82-87; 120pp; English.

XX The invention relates to a composition comprising ribavirin and an

CC antigen preferably non structural 3 protein (NS3)/4A fragment of

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV

CC sequence. The composition is useful for enhancing an immune response to

CC a hepatitis C antigen in humans, domestic, sport or pet species and as

CC vaccines for treating and preventing HCV infections. The composition is

CC also useful for treating viral, bacterial, fungal diseases and cancer.

CC The present sequence is hepatitis A virus (HAV) protein.

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20
 |||||
 931 DLEETIANSKDFPNMSETDL 950

RESULT 13

ABU08639
 ID ABU08639 standard; Protein; 2227 AA.

AC ABU08639;

DT 03-JUN-2003 (first entry)

DE Wild type human hepatitis A virus strain HM-175.

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KV vaccine; MRC-5 cell; hepatitis infection.

OS Hepatitis A virus strain HM-175.

PN US2002176869-A1.

PD 28-NOV-2002.

PF 29-APR-2002; 2002US-0135988.

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-0508610.

PR 17-APR-1995; 95US-0397232.

PA (FUNK/) FUNKHOUSER A W.

PA (EMER/) EMERSON S U.

PA (PURC/) PURCELL R H.

PA (DHON/) D'HONDT E.

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

PI MPI; 2003-352605/33.

DR N-PSDB; ABX93473.

PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,

PT and disease -

XX The invention describes a live hepatitis A virus (HAV) adapted to growth

XX in MRC-5 cells. The HAV and compositions comprising the HAV are useful as

XX a vaccine for protecting primates against hepatitis infection and

XX disease. This is the amino acid sequence of wild type human hepatitis A

XX virus strain HM-175.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 24; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
 |||||
 Db 931 DLEETIANSKDFPNMSETDL 950

RESULT 14

ABU08640
 ID ABU08640 standard; Protein; 2227 AA.

AC ABU08640;

DT 03-JUN-2003 (first entry)

DE Attenuated (pass35) hepatitis A virus strain HM-175.

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KV vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.

OS Hepatitis A virus strain HM-175.

PN US2002176869-A1.

PD 28-NOV-2002.

PF 29-APR-2002; 2002US-0135988.

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-0508610.

PR 17-APR-1995; 95US-0397232.

PA (FUNK/) FUNKHOUSER A W.

PA (EMER/) EMERSON S U.

PA (PURC/) PURCELL R H.

PA (DHON/) D'HONDT E.

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

PI MPI; 2003-352605/33.

DR N-PSDB; ABX93474.

PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,

PT and disease -

XX The invention describes a live hepatitis A virus (HAV) adapted to growth

XX in MRC-5 cells. The HAV and compositions comprising the HAV are useful as

XX a vaccine for protecting primates against hepatitis infection and

XX disease. This is the amino acid sequence of an attenuated (pass 35)

XX human hepatitis A virus strain HM-175.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 24; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
 |||||
 Db 931 DLEETIANSKDFPNMSETDL 950

RESULT 15

ABU08641

ID	AB008641	standard; Protein; 2227 AA.
XX		
AC	AB008641;	
XX		
DT	03-JUN-2003	(first entry)
XX		
DE	Attenuated hepatitis A virus (4380) strain HM-175.	
XX		
KV	Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;	
KW	vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.	
XX		
OS	Hepatitis A virus strain HM-175.	
XX		
PN	US2002176869-A1.	
PD		
XX		
XX	28-NOV-2002.	
XX		
PE	29-APR-2002; 2002US-0135988.	
XX		
XX	07-JUN-1995; 95US-0475886.	
XX	31-AUG-2000; 2000US-0653499.	
XX	18-SEP-1992; 92US-0947338.	
PR	17-SEP-1993; 93WO-US08610.	
PR	17-APR-1995; 95US-0397232.	
XX		
PA	(FUNK/) FUNKHOUSER A W.	
PA	(EMER/) EMERSON S U.	
PA	(PURC/) PORCELL R H.	
PA	(DHON/) D'HONDT E.	
XX		
PI	Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;	
XX		
DR	WPI: 2003-352605/33.	
XX	N-PSDB; ABX93475.	
XX		
PT	New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,	
PT	useful in vaccines for protecting primates against hepatitis infection	
PT	and disease	
XX		
PS	Disclosure: Page 45-51; 70pp; English.	
XX		
CC	The invention describes a live hepatitis A virus (HAV) adapted to growth	
CC	in MRC-5 cells. The HAV and compositions comprising the HAV are useful as	
CC	a vaccine for protecting primates against hepatitis infection and	
CC	disease. This is the amino acid sequence of an attenuated human	
CC	hepatitis A virus (4380) strain HM-175.	
XX		
SO	Sequence 2227 AA;	
XX		
OY	Very Match	100.0%; Score 101; DB 24; Length 2227;
DB	Best Local Similarity	100.0%; Pred. No. 2.5e-07;
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	1 DLEIIAANSKDFPMWSTDL 20	
	931 DLEIIAANSKDFPMWSTDL 950	

```
Search completed: October 1, 2003, 09:56:44
Job time : 44.5294 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 : Search time 12.8235 Seconds
(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101
Sequence: 1 DLEEIANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	1 GNNYHM	genome polyprotein
2	101	100.0	2227	1 GNNYHR	genome polyprotein
3	101	100.0	2227	1 GNNYMK	genome polyprotein
4	101	100.0	2227	1 GNNYHB	genome polyprotein
5	98	97.0	2230	1 GNNYSA	genome polyprotein
6	50	49.5	736	2 D90574	hypothetical prote
7	49	48.5	400	2 F88931	protein R1G11.1 (
8	45.5	45.0	930	2 A84668	Argonaute (AGO)-1
9	45.5	45.0	1621	2 T15264	hypothetical prote
10	44	43.6	740	1 FOLJHD	gag polyprotein -
11	44	43.6	927	2 T38127	phosphoprotein - f
12	44	43.6	1046	2 S67786	hypothetical prote
13	43.5	43.1	754	1 BABOH	peptide-aspartate
14	43.5	43.1	1451	2 S65571	pattern formation
15	43.5	43.1	5107	2 T29144	partial CDS - Caen
16	43	42.6	152	2 A64610	hypothetical prote
17	43	42.6	152	2 A71904	hypothetical prote
18	43	42.6	231	2 B70407	probable di-trans,
19	43	42.6	285	2 B87426	rhodanese family p
20	43	42.6	454	2 T02100	hypothetical prote
21	43	42.6	598	2 H71336	probable cell divi
22	43	42.6	1072	2 A84112	alkaline amylopull
23	43	42.6	1201	2 T08603	kinasin-related pr
24	42.5	42.1	913	2 T15278	hypothetical prote
25	42.5	42.1	4717	2 T41581	hypothetical colle
26	42	41.6	289	2 T23342	hypothetical prote
27	42	41.6	365	2 C81050	cytochrome c oxida
28	42	41.6	365	2 F81826	probable cytochrom
29	42	41.6	538	2 G95015	ABC transporter, A

30	42	41.6	538	2 A97889	hypothetical prote
31	42	41.6	1112	2 T47784	hypothetical prote
32	42	41.6	1946	2 AE1449	hypothetical prote
33	41.5	41.1	236	2 A81283	probable periplasm
34	41.5	41.1	694	2 S41868	DNA-directed RNA p
35	41.5	41.1	1378	2 A81393	hypothetical prote
36	41	40.6	94	2 D70245	hypothetical prote
37	41	40.6	108	2 AF2042	hypothetical prote
38	41	40.6	133	2 A29174	clostridin (EC 3.
39	41	40.6	139	2 T26224	hypothetical prote
40	41	40.6	230	2 T45754	hypothetical prote
41	41	40.6	262	2 E89760	conserved hypotet
42	41	40.6	292	2 F86431	hypothetical prote
43	41	40.6	295	2 H83642	probable 2-OH-laur
44	41	40.6	346	2 G64182	lipopolysaccharide
45	41	40.6	368	2 G97291	hypothetical prote

ALIGNMENTS

RESULT 1

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
J:Coheh, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J:Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with
A:Reference number: A25981; MID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <CON>
A:Cross-references: EMBL:M14707; NID:9329582; PIDN:AAA45465.1; PID:9329583
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr
F:1-23/Product: coat protein 1A #status predicted <VP1>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP4>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1432/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIANSKDFPNMSETDL 20
Db 931 DLEEIANSKDFPNMSETDL 950

RESULT 2
GNNYHR
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:R02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match
Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 931 DLEIANSKDFPMSETDL 20
1 DLEIANSKDFPMSETDL 20
|||||

RESULT 3
GNNYMK
genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence-revision 30-Jun-1988 #text-change 16-Jul-1999
C:Accession: A94149; A259149; A34508
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Felstone, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with reference number: A94149; MUID:8715701; PMID:3031686
A:Accession: A94149
A:Reference number: A94149; MUID:8715701; PMID:3031686
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: submitted to Genbank, August 1987
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: coat protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match
Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 931 DLEIANSKDFPMSETDL 20
1 DLEIANSKDFPMSETDL 20
|||||

RESULT 4
GNNYHB
genome polypeptide - human hepatitis A virus (strain MB8)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence-revision 30-Jun-1991 #text-change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klein, R.; Wimmer, E.; Delnath
virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-461/Product: coat protein 1C #status predicted <VP3>
F:462-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match
Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 931 DLEIANSKDFPMSETDL 20
1 DLEIANSKDFPMSETDL 20
|||||

RESULT 5
GNNYSA
genome polypeptide - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence-revision 31-Mar-1992 #text-change 16-Jun-2000
R:Tsarev, S.A.
submitted to JFPI, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:DO0924; NID:g222597; PIDN:BA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapalidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapalidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-27/Product: coat protein 1A #status predicted <C1A>

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence-revision 30-Jun-1991 #text-change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klein, R.; Wimmer, E.; Delnath
virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-461/Product: coat protein 1C #status predicted <VP3>
F:462-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match
Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 931 DLEIANSKDFPMSETDL 20
1 DLEIANSKDFPMSETDL 20
|||||

RESULT 5
GNNYSA
genome polypeptide - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 3A; core protein 3B; core protein 3C; core protein 3D
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence-revision 31-Mar-1992 #text-change 16-Jun-2000
R:Tsarev, S.A.
submitted to JFPI, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:DO0924; NID:g222597; PIDN:BA00766.1; PID:g222598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: J01080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjapalidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA33490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjapalidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-495/Product: coat protein 1C #status predicted <C1C>
 F:496-795/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.0% Score 98; DB 1; Length 2230;
 Best Local Similarity 95.0%; Pred. No. 6e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNMSETDL 20
 |||||:|||||:|||||:
 Db 935 DLEIANSKDFPNMSETDL 954

ULT 6
 0574

hypothetical protein MYPV_5000 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
 C:Species: Mycoplasma pulmonis
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
 C:Accession: D90574
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, I.;
 Nucleic Acids Res. 29, 2145-2153, 2001
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
 A:Reference number: A9512; MUID:21267165; PMID:11353084
 A:Accession: D90574
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-736 <KOR>
 A:Cross-references: GB:AL445566; PID:q14089914; PIDN:CAC13673.1; GSPDB:GN00153
 A:Experimental source: strain UAB CTIP
 C:Genetics:
 A:Gene: MYPV_5000
 A:Genetic code: SGC3

Query Match 49.5% Score 50; DB 2; Length 736;
 Best Local Similarity 55.0%; Pred. No. 8.3;
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNMSETDL 20
 |||||:|||||:|||||:
 Db 218 DLYSIYNEKDAPEISEEDL 237

SULF 7
 F88931

protein RL1611.1 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: F88931
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele
 A:Note: published extra appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A:Accession: F88931
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-400 <STO>
 A:Cross-references: GB:chr_V; PIDN:AAC69076.1; PID:q2384851; GSPDB:GN00023; CESP:RL1611.
 C:Genetics:
 A:Gene: RL1611.1
 A:Map position: 5

Query Match 48.5% Score 49; DB 2; Length 400;
 Best Local Similarity 47.4%; Pred. No. 5.9;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 OY 2 LEEIANSKDFPNMSETDL 20
 |||||:|||||:|||||:
 Db 164 LENTSTNMKDFHSSKEDV 182

RESULT 8
 A84668

Argonate (AGO1)-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: A84668
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: A84668
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-930 <STO>
 A:Cross-references: GB:AE002093; NID:g3885334; PIDN:AAC77862.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g27040
 A:Map position: 2
 C:Superfamily: rabbit translation initiation factor eIF-2C

Query Match 45.0% Score 45.5; DB 2; Length 930;
 Best Local Similarity 42.3%; Pred. No. 57;
 Matches 11; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

OY 2 LEEIANSKDFPNMSETDL 18
 |||||:|||||:|||||:
 Db 153 LEEVATSKDFVSRANGSPNGNS 178

RESULT 9
 T15264

hypothetical protein F59E12.9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T15264
 R:Johnson, D.
 submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid F59E12.
 A:Reference number: Z18318
 A:Accession: T15264
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1621 <JOH>
 A:Cross-references: EMBL:AF003366; NID:g2088833; PID:g2088843; PIDN:AAB54259.1; GSPDB
 A:Experimental source: strain Bristol N2; clone F59E12
 C:Genetics:
 A:Gene: CESP:F59E12.9
 A:Map position: 2
 A:Introns: 30/3; 55/1; 200/2; 289/2; 327/2; 369/3; 589/3; 860/1; 1278/1; 1547/

Query Match 45.0% Score 45.5; DB 2; Length 1621;
 Best Local Similarity 42.9%; Pred. No. 1.1e+02;
 Matches 9; Conservative 8; Mismatches 1; Indels 3; Gaps 1;

OY 2 LEEIANSKDFPNMSETDL 19
 |||||:|||||:|||||:
 Db 985 IKEIYASAKDFMDDVSDSD 1005

RESULT 10
 FOLJHD

gag polyprotein - squirrel monkey retrovirus SMRV-H
 N:Contains: core protein p16; core protein p19; probable core protein p10; probable c
 C:Species: squirrel monkey retrovirus SMRV-H

C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
 C/Accession: A31827
 R/Oda, T.; Ikeda, S.; Matsushika, M.; Akiyama, K.; Mitsunobu, F.
 Virology 167, 468-476, 1988
 A>Title: Molecular cloning, complete nucleotide sequence, and gene structure of the pro
 A/Reference number: A31827; MUID:9073750; PMID:3201749
 A/Accession: A31827
 A/Molecule type: DNA
 A/Residues: 1-740 <ODA>
 A/Cross-references: GB:M23385; NID:g332626; PIDN:AAA6451.1; PID:g807672
 C/Genetics:
 A:Gene: gag
 C/Superfamily: AIDS-related virus gag polyprotein
 C/Keywords: core protein, polyprotein
 F:1163/Product: core protein p19 #status predicted <CP9>
 F:164-318/Product: core protein p16 #status predicted <CP6>
 F:319-648/Product: core protein p35 #status predicted <CP5>
 F:649-740/Product: core protein p10 #status predicted <CP1>

Query Match 43.6%; Score 44; DB 1; Length 740;
 Best Local Similarity 50.0%; Pred. No. 76;
 Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

OY 1 DLEBIAA-NSKDFPNMSET 18
 Db 214 DLEBAAQYNNPDPQLTNT 233

RESULT 11

T38127

phosphoprotein - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Aug-2002

C/Accession: T38127; A57087; S54119

R/Baddock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, May 1997

A/Reference number: Z21772

A/Accession: T38127

A/Status: preliminary; translated from GB/EMBL/DBDUT

A/Molecule type: DNA

A/Residues: 1-927 <BAD>

A/Cross-references: EMBL:Z95334; PIDN:CAB08599.2; GSPDB:GN00066; SPDB:SPAC2068.05C

A/Experimental source: strain 9729; cosmid c20G8

R/Fankhauser, C.; Raymond, A.; Cerutti, L.; Utzig, S.; Hofmann, K.; Simanis, V.
 Cell 82, 435-444, 1995

A>Title: The Schizosaccharomyces pombe cdc15 gene is a key element in the reorganization

A/Reference number: A57087; MUID:95360987; PMID:7634333

A/Accession: A57087

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: MLTKSLQ', 28-927 <FAN>

A/Cross-references: GB:X86179

R/Fankhauser, C.; Raymond, A.; Cerutti, L.; Simanis, V.
 submitted to the EMBL Data Library, April 1995

A/Description: The cdc15 gene is a key element in F-actin reorganisation at mitosis.

A/Reference number: S54119

A/Accession: S54119

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 'MLTKSLQ', 28-870, 'GKFIKEL' <FAZ>

A/Cross-references: EMBL:X86179

C/Genetics:

A:Gene: SPDB:SPAC2068.05c; cdc15

A/Map position: 1

A/Introns: 27/3; 58/2; 871/1

C/Superfamily: fission yeast scd2 protein; SH3 homology

C/Keywords: mitosis; phosphoprotein

F:873-924/Domain: SH3 homology <SH3>

Query Match 43.6%; Score 44; DB 2; Length 927;
 Best Local Similarity 47.1%; Pred. No. 98;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 LEEIANSKDFPNMSET 18
 Db 69 LOELIASSADIPVGS 85

RESULT 12

S67786

hypothetical protein YDU223c - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein D0843

C/Species: Saccharomyces cerevisiae

C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002

C/Accession: S67786

R/Rasmussen, S.W.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S67778

A/Accession: S67786

A/Molecule type: DNA

A/Residues: 1-1046 <RAS>

A/Cross-references: EMBL:Z74271; NID:g1431374; PID:e253363; PID:g1431375; GSPDB:GN000

A/Experimental source: strain S288C

C/Genetics:

A:Gene: MIPS:YDL223c

A/Cross-references: SGD:S0002382

A/Map position: 4L

Query Match 43.6%; Score 44; DB 2; Length 1046;
 Best Local Similarity 50.0%; Pred. No. 1,1e+02;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 EELIANSKDFPNMSET 18
 Db 733 QDIASDAKDFNNPET 748

RESULT 13

BAB0H

peptide-aspartate beta-dioxygenase (EC 1.14.11.16) - bovine

N/Alternate names: aspartyl (asparaginyl) beta-hydroxylase

C/Species: Bos primigenius taurus (cattle)

C/Date: 31-Dec-1993 #sequence_revision 10-Feb-1995 #text_change 11-Jun-1999

C/Accession: A42969; A39470; B39470; C39470; S27948

R/Jia, S.; Vandusen, W.J.; Diehl, R.E.; Kohl, N.E.; Dixon, R.A.; Elliston, K.O.; Ster

J. Biol. Chem. 267, 14322-14327, 1992

A>Title: cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-hydroxylas

A/Reference number: A42969; MUID:92332546; PMID:1378441

A/Accession: A42969

A/Molecule type: mRNA

A/Residues: 1-754 <TIA>

A/Cross-references: EMBL:M91213; NID:g162693; PIDN:AAA03563.1; PID:g162694

A/Experimental source: brain

A/Note: Sequence extracted from NCBI backbone (NCBI:108534)

R/Wang, O.; Vandusen, W.J.; Petroski, C.J.; Garsky, V.M.; Stern, A.M.; Friedman, P.A.
 J. Biol. Chem. 266, 14004-14010, 1991

A>Title: Bovine liver aspartyl beta-hydroxylase. Purification and characterization.

A/Reference number: A39470; MUID:91310689; PMID:1856229

A/Accession: A39470

A/Molecule type: protein

A/Residues: 615, 'X', 617-630, 'XX', 633-634, 'X', 636, 'XX', 639-641 <MA2>

A/Accession: C39470

A/Molecule type: protein

A/Residues: 311-347, 'X', 349, 'X', 351-373, 'X', 375-379, 'X', 381-382 <MA3>

C/Comment: This enzyme uses ferrous iron as a cofactor, and while beta-hydroxylating

C/Comment: Aspartic acid and asparagine residues in the BGF homology domain of certai

C/Superfamily: peptidyl-aspartate beta-dioxygenase; tetrairicopeptide repeat homology

C/Keywords: glycoprotein; oxidoreductase; transmembrane protein

F:2-56/Domain: Intracellular #status predicted <IRM>

F:57-78/Domain: transmembrane #status predicted <IRM>

F:289-754/Product: peptidyl-aspartate beta-dioxygenase, 56K form #status predicted <56

F:311-354/Product: peptidyl-aspartate beta-dioxygenase, 52K form #status predicted <52

F:337-370/Domain: tetrairicopeptide repeat homology <TT1>

F:371-404/Domain: tetratricopeptide repeat homology <TRP>
F:13,96,466,702/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 43.1%; Score 43.5; DB 1; Length 754;
Best Local Similarity 55.6%; Pred. No. 93;

Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

OY 1 DLEETIANSKDFP-NMSE 17
DB 82 DYEETIANKAKDFRYNLSE 99

RESULT 14

S65571

Pattern formation protein GNOM - Arabidopsis thaliana

N:Alternate names: EMB30 protein

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 07-May-1999

C:Accession: S65571; S65572

C:Author: Busch, M.; Mayer, U.; Jurgens, G.

C:Title: Molecular analysis of the Arabidopsis pattern formation gene GNOM: gene structure

A:Reference number: S65571; MUID:96204508; PMID:8628228

A:Accession: S65571

A:Molecule type: DNA

A:Residues: 1-1451 <BUS>

A:Cross-References: EMBL:U36433; NID:g1209632; PID:g1209633

A:Accession: S65572

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-110, 'T', 112-866, 'G', 868-1451 <BMW>

A:Cross-References: EMBL:U36432; NID:g1209630; PID:g1209631

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995

C:Genetics:

A:Gene: GNOM: EMB30

A:introns: 246/3

Query Match 43.1%; Score 43.5; DB 2; Length 1451;
Best Local Similarity 63.2%; Pred. No. 2e+02;

Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 1 DLEETIANS-KDFPNNMSE 18
DB 1302 DLEETIAGSQKDYRNMEGT 1320

RESULT 15

T29144

Partial CDS - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000

C:Accession: T29144

C:Author: R. Pauley, A.; Galtung, S.

C:Submitted to the EMBL Data Library, July 1996

A:Description: The sequence of C. elegans cosmid K11C4.

A:Reference number: Z20577

A:Accession: T29144

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5107 <PAD>

A:Cross-References: EMBL:U64854; PIDN:AA18318.1; GSPDB:GN00023; CESP:unc-68

A:Experimental source: strain Bristol N2; clone K11C4

C:Genetics:

A:Gene: CESP:unc-68

A:Map position: 5

A:introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/

/3; 3369/2; 3313/3; 3466/1; 3519/3; 3629/3; 3658/2; 3710/1; 3741/3; 3779/2; 3810

C:Superfamily: Ryanodine receptor; transcription initiation factor sigma region 1 homolog

Query Match 43.1%; Score 43.5; DB 2; Length 5107;
Best Local Similarity 43.5%; Pred. No. 8 4e+02;

Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

OY 1 DLEETIANS---KDFPNNMSETDL 20
DB 3377 DLEETIANNNTMTSDVPNVYDVL 3399

Search completed: October 1, 2003, 10:04:43
Job time : 15.8235 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds
(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-45
Perfect score: 101
Sequence: 1 DLEETAAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2226	1	P26581 hepatitis a
2	101	100.0	2226	1	P26582 hepatitis a
3	101	100.0	2227	1	P08617 hepatitis a
4	101	100.0	2227	1	P06441 hepatitis a
5	101	100.0	2227	1	P13501 hepatitis a
6	98	97.0	2230	1	P14553 simian hepa
7	97	96.0	2226	1	P26580 hepatitis a
8	46	45.5	194	1	P57674 thermoplasma
9	44	43.6	740	1	P21311 squirrel mo
10	44	43.6	927	1	P09822 schistosom
11	43.5	43.1	754	1	P26581 hepatitis a
12	43.5	43.1	1451	1	P08617 hepatitis a
13	43	42.6	231	1	P13501 hepatitis a
14	43	42.6	350	1	P14553 simian hepa
15	41.5	41.1	1378	1	P26580 hepatitis a
16	41	40.6	346	1	P57674 thermoplasma
17	41	40.6	376	1	P21311 squirrel mo
18	41	40.6	526	1	P09822 schistosom
19	41	40.6	528	1	P26581 hepatitis a
20	41	40.6	548	1	P08617 hepatitis a
21	40	39.6	283	1	P13501 hepatitis a
22	40	39.6	331	1	P14553 simian hepa
23	40	39.6	381	1	P26580 hepatitis a
24	40	39.6	547	1	P57674 thermoplasma
25	40	39.6	824	1	P21311 squirrel mo
26	40	39.6	1087	1	P09822 schistosom
27	39.5	39.1	244	1	P26581 hepatitis a
28	39.5	39.1	521	1	P08617 hepatitis a
29	39.5	39.1	552	1	P13501 hepatitis a
30	39	38.6	221	1	P14553 simian hepa
31	39	38.6	223	1	P26580 hepatitis a
32	39	38.6	253	1	P57674 thermoplasma
33	39	38.6	320	1	P21311 squirrel mo

34	39	38.6	394	1	UXNA_SALTY
35	39	38.6	396	1	LDLD_ECOLI
36	39	38.6	396	1	YJEB_YEAST
37	39	38.6	400	1	DDX1_DROVI
38	39	38.6	438	1	GUX3_AGABI
39	39	38.6	467	1	MURD_BRUME
40	39	38.6	467	1	MURD_BRUSU
41	39	38.6	474	1	CC2H_CRIFA
42	39	38.6	631	1	PAB3_HUMAN
43	39	38.6	642	1	YG22_YEAST
44	39	38.6	660	1	DNAR_CHLPN
45	39	38.6	673	1	VID3_AGRTS

ALIGNMENTS

RESULT 1
POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID POLG_HPAV4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae; Hepatovirus.
OC NCBI_TaxID=12095;
OX [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Reinstone S.M., Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination".
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59809; AAA5469.1; -
CC MEROPS: C03.005; -
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polyprotein; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1
FT CHAIN 23
FT CHAIN 24
FT CHAIN 245
FT CHAIN 246
FT CHAIN 491
FT CHAIN 492
FT CHAIN 794
FT CHAIN 795
FT CHAIN 900
FT CHAIN 901
FT CHAIN 1087
FT CHAIN 1088
FT CHAIN 1422
CORE PROTEIN P2B.
CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BE75 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEFIANSKDFPNMSETDL 20
DB 931 DLEFIANSKDFPNMSETDL 950

RESULT 2
POLG_HPAV8 STANDARD: PRT: 2226 AA.
ID AC P26582; 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Ciomeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59808: AAA45467.1; -
CC PDB: 1OAT; 15-MAY-00.
CC MEROPS: C03.005; -
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC Polyprotein; Coat protein; Core protein; Thiol protease;
CC RNA-directed RNA polymerase; Hydrolyase; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
COAT PROTEIN VP2 (P1A).
COAT PROTEIN VP3 (P1C).
COAT PROTEIN VP1 (P1D).
CORE PROTEIN P2A.
CORE PROTEIN P2B.
CORE PROTEIN P2C.
PROBABLE PROTEIN P3A.

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396CB06B CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEFIANSKDFPNMSETDL 20
DB 931 DLEFIANSKDFPNMSETDL 950

RESULT 3
POLG_HPAV8 STANDARD: PRT: 2227 AA.
ID AC P08617; 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DR 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59(1987).
CC [2]
CC SEQUENCE FROM N.A.
CC RC STRAIN-Attenuated;
CC RX MEDLINE=87175701; PubMed=3031686;
CC Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
CC Purcell R.H.;
CC "Complete nucleotide sequence of an attenuated hepatitis A virus:
CC comparison with wild-type virus.";
CC Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
CC [3]
CC SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
CC MEDLINE=85166289; PubMed=2984684;
CC Baroudy B.M., Ticehurst J.R., Miele T.A., Malzel J.V. Jr.,
CC Purcell R.H., Feinstone S.M.;
CC "Sequence analysis of hepatitis A virus cDNA coding for capsid
CC proteins and RNA polymerase.";
CC Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@lsb.slb.ch).

CC EMBL: M14114; AAA45475.1; -
 DR EMBL: M14707; AAA45465.1; -
 DR EMBL: M14707; AAA45466.1; ALT_INIT.
 DR EMBL: M16632; AAA45471.1; -
 DR PIR: A25981; GNNYMK.
 DR PIR: A94149; GNNYMK.
 DR MEROPS: C03.005; -
 DR PDB: 1HAV; 23-DEC-96.
 DR INTERPRO: IPR004004; Calic1.pol_hel.
 DR INTERPRO: IPR000605; RNA_helicase.
 DR INTERPRO: IPR007095; RNA_pol_DS_PS.
 DR INTERPRO: IPR001205; RNA_pol_P3D.
 DR INTERPRO: IPR007094; RNA_pol_PSVlr.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 KM Polypeptide; Coat protein; Core protein; Transferase;
 RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 CHAIN 1 23
 FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
 FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
 FT CHAIN 482 836 COAT PROTEIN VP3 (PIC).
 FT CHAIN 837 980 COAT PROTEIN VP1 (PID).
 FT CHAIN 981 1087 CORE PROTEIN P2A.
 FT CHAIN 1088 1422 CORE PROTEIN P2B.
 FT CHAIN 1423 1496 CORE PROTEIN P2C.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3A.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3B.
 FT CHAIN 1739 2227 PROBABLE PROTEIN P3C.
 FT CHAIN 2227 2227 RNA-DIRECTED POLYMERASE 3D.
 FT VARIANT 77 77 K -> R (IN ATTENUATED STRAIN).
 FT VARIANT 764 764 E -> V (IN ATTENUATED STRAIN).
 FT VARIANT 821 821 N -> S (IN ATTENUATED STRAIN).
 FT VARIANT 1052 1052 A -> V (IN ATTENUATED STRAIN).
 FT VARIANT 1062 1062 G -> A (IN ATTENUATED STRAIN).
 FT VARIANT 1118 1118 K -> M (IN ATTENUATED STRAIN).
 FT VARIANT 1151 1151 E -> S (IN ATTENUATED STRAIN).
 FT VARIANT 1163 1163 F -> S (IN ATTENUATED STRAIN).
 FT VARIANT 1277 1277 H -> I (IN ATTENUATED STRAIN).
 FT VARIANT 1500 1500 V -> Y (IN ATTENUATED STRAIN).
 FT VARIANT 1805 1805 D -> N (IN ATTENUATED STRAIN).
 FT VARIANT 1930 1930 S -> T (IN ATTENUATED STRAIN).
 SO SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20
 DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 4
 POLG_HPAVL STANDARD; PRT; 2227 AA.
 ID POLG_HPAVL
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 41, Last annotation update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain LA).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC Hepatovirus.
 OC NCBI_TaxID=12099;
 OX NCBI_TaxID=12099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., Van Nest G., Dina D.;

Primary structure and gene organization of human hepatitis A virus.;
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICSAHEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.lsb.slb.ch/announce/>
 CC or send an email to license@lsb.slb.ch).

CC EMBL: K02990; AAA45472.1; -
 DR PIR: A03903; GNNYMK.
 DR MEROPS: C03.005; -
 DR INTERPRO: IPR004004; Calic1.pol_hel.
 DR INTERPRO: IPR000605; RNA_helicase.
 DR INTERPRO: IPR007095; RNA_pol_DS_PS.
 DR INTERPRO: IPR001205; RNA_pol_P3D.
 DR INTERPRO: IPR007094; RNA_pol_PSVlr.
 DR Pfam: PF00680; RNA_dep_RNA_pol.1.
 DR Pfam: PF00910; RNA_helicase.1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 KM Polypeptide; Coat protein; Core protein; Transferase;
 RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 CHAIN 1 23
 FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
 FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
 FT CHAIN 492 836 COAT PROTEIN VP3 (PIC).
 FT CHAIN 837 980 COAT PROTEIN VP1 (PID).
 FT CHAIN 981 1076 CORE PROTEIN P2A.
 FT CHAIN 1077 1422 CORE PROTEIN P2B.
 FT CHAIN 1423 1484 CORE PROTEIN P2C.
 FT CHAIN 1485 1507 PROBABLE PROTEIN P3A.
 FT CHAIN 1508 1678 PROBABLE PROTEIN P3B.
 FT CHAIN 1679 2227 PROBABLE PROTEIN P3C.
 SO SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20
 DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 5
 POLG_HPAVM STANDARD; PRT; 2227 AA.
 ID POLG_HPAVM
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain HMB).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC Hepatovirus.
 OC NCBI_TaxID=12100;
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Klehn R., Wimmer E., Deinhardt F.;

RT "The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MB)."
RL Virus Res. 8:153-171(1987).
CC
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA](N).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC
CC -1- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: M20273; AAA45474.1; -
CC
CC MEROPS: C03.005; -
CC
CC InterPro: IPR000605; RNA_helicase.
CC
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC
CC InterPro: IPR001205; RNA_pol_P3D.
CC
CC InterPro: IPR007094; RNA_pol_PSVir.
CC
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC
CC Pfam: PF00910; RNA_helicase; 1.
CC
CC Polyprotein: Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CC
CC CHAIN 1 23
CC FT CHAIN 24 245
CC FT CHAIN 246 491
CC FT CHAIN 492 836
CC FT CHAIN 837 980
CC FT CHAIN 981 1087
CC FT CHAIN 1088 1422
CC FT CHAIN 1423 1496
CC FT CHAIN 1497 1519
CC FT CHAIN 1520 1738
CC FT CHAIN 1739 2227
CC SO SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1;le-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 DLEBIANSKDFPMSETDL 20
931 DLEBIANSKDFPMSETDL 950
DLEBIANSKDFPMSETDL 950

RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RX SEQUENCE OF 1750-2164 FROM N.A.
RA MEDLINE-69232168; PubMed-2541023;
RA Balayan M.S., Kosov Y.T., Andjapartidze A.G., Tsarev S.A.,
RA Tsarev S.A., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA](N).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC
CC -1- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: D00924; BA00766.1; -
CC
CC EMBL: X15461; CA033490.1; -
CC
CC PIR: A30470; GNNYSA.
CC
CC MEROPS: C03.005; -
CC
CC InterPro: IPR004004; Calic1_pol_hel.
CC
CC InterPro: IPR000605; RNA_helicase.
CC
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC
CC InterPro: IPR001205; RNA_pol_P3D.
CC
CC InterPro: IPR007094; RNA_pol_PSVir.
CC
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC
CC Pfam: PF00910; RNA_helicase; 1.
CC
CC PRINTS: PR00918; CALICIVIRUSNS.
CC
CC Polyprotein: Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CC
CC CHAIN 1 27
CC FT CHAIN 28 249
CC FT CHAIN 250 495
CC FT CHAIN 496 795
CC FT CHAIN 796 984
CC FT CHAIN 985 1091
CC FT CHAIN 1092 1426
CC FT CHAIN 1427 1498
CC FT CHAIN 1499 1521
CC FT CHAIN 1522 1741
CC FT CHAIN 1742 2230
CC SO SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
Query Match 97.0%; Score 98; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 3;3e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 DLEBIANSKDFPMSETDL 20
935 DLEBIANSKDFPMSETDL 954
DLEBIANSKDFPMSETDL 954


```

DR PIR: A31827; FOLJHD.
DR InterPro: IPR003322; Gag_P10.
DR InterPro: IPR000721; Gag_P24.
DR InterPro: IPR001878; Znf_CCHC.
DR Pfam: PF02337; Gag_P10; 1.
DR Pfam: PF00607; Gag_P24; 1.
DR Pfam: PF00098; zfc_CCHC; 1.
DR SMART: SM00343; Znf_C2HC; 1.
DR Core protein; Polyprotein.
FT CHAIN 1 163 CORE PROTEIN P19.
FT CHAIN 164 318 CORE PROTEIN P16.
FT CHAIN 319 648 PROBABLE CORE PROTEIN P35.
FT CHAIN 649 740 PROBABLE CORE PROTEIN P10.
SQ SEQUENCE 740 AA; 80543 MW; CC2503C2661221F5 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 740;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;

1 DLELIAA-NSKDPNNSET 18
||||| : : : :
214 DLEIAAAYNNPWPOLJNT 233

RESULT 10
CC15_SCHPO STANDARD; PRT; 927 AA.
AC Q09822; O14365; Q90U50;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-SEP-2003 (Rel. 39, Last sequence update)
DE Cell division control protein 15.
GN CDC15 OR SPAC2068.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC NCBI_TaxID=4896;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-95360987; PubMed-7634333;
RA Fankhauser C., Raymond A., Cerutti L., Utzig S., Hofmann K.,
RA Simanis V.;
RT "the S. pombe cdc15 gene is a key element in the reorganization of F-
actin at mitosis.";
RL Cell 82:435-444(1995).
RN [2]
REVISONS TO N-TERMINUS.
Fankhauser C., Raymond A., Cerutti L., Utzig S., Hofmann K.,
Simanis V.;
Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-972;
RX MEDLINE-21848401; PubMed-11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris A., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voicakeert G., Aert R., Robben J., Glynn P.,
RA Wellens I., Vanstreets E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,

```

```

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lalaurie V., Mottier S.,
RA Gallibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Palsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [4]
RP SEQUENCE OF 286-425 FROM N.A.
RC STRAIN-968 h90;
RX MEDLINE-20223868; PubMed-10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Harauchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
CC -1- FUNCTION: AFTER THE ONSET OF MITOSIS, FORMS A RING-LIKE STRUCTURE
WHICH CO-LOCALIZES WITH THE MEDIAL ACTIN RING. APPEARS TO MEDATE
CYTOSKELETAL REARRANGEMENTS REQUIRED FOR CYTOKINESIS. ESSENTIAL
FOR VIABILITY.
CC -1- DEVELOPMENTAL STAGE: PEAKS IN EARLY MITOSIS BEFORE SEPARATION.
CC -1- DOMAIN: THE N-TERMINAL REGION IS IN A COILED COIL STRUCTURE.
CC -1- PTM: PHOSPHORYLATED.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -1- SIMILARITY: Contains 1 SH3 domain.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X86179; CA60115.1; -
DR EMBL: Z95334; CAB08599.2; -
DR EMBL: AB027810; BAA87114.1; -
DR PIR: T38127; T38127.
DR HSSP: P07751; TRUD.
DR GenBank: spombe; SPAC2068.05C; -
DR InterPro: IPR001060; CC15_Fes_CTP4.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001060; CC15_Fes_CTP4.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00611; FCH; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR SMART: SM00055; SH3; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50133; FCH; 1.
DR PROSITE: PS50002; SH3; 1.
DR MitoSis; Cytoskeleton; SH3 domain; Phosphorylation; Coiled coil.
FT DOMAIN 20 101 FCH.
FT DOMAIN 108 207 COILED COIL (POTENTIAL).
FT DOMAIN 866 927 SH3.
SQ SEQUENCE 927 AA; 102119 MW; FDCE7ED0AA3D247D CRC64;

Query Match 43.6%; Score 44; DB 1; Length 927;
Best Local Similarity 47.1%; Pred. No. 43;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

2 LLEIAANSKDPNNSET 18
|:|:|:|:|:|
Db 69 LQELIAASSADIPWGST 85

RESULT 11
ASPH_BOVIN STANDARD; PRT; 754 AA.
ID ASPH_BOVIN
AC Q28056;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

```

01-NOV-1997 (Rel. 35, Last annotation update)
DE ASPARTYL/asparaginyl beta-hydroxylase (EC 1.14.11.16) (aspartate beta-
DE hydroxylase) (ASP beta-hydroxylase) (Peptide aspartate beta-
DE dihydroxylase).
GN ASPH.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bos.
OX NCBI_taxid:9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Liver;
RX MEDLINE=92332546; PubMed=1378441;
RA Jia S., Vandenusen W.J., Diehl R.E., Kohl N.E., Dixon R.A.F.,
RT Elliston K.O., Stern A.M., Friedman P.A.;
RT "cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-
RT hydroxylase";
RL J. Biol. Chem. 267:14322-14327(1992).
[2]
SEQUENCE OF 289-385 AND 615-641.
TTISSUE=Liver;
RX MEDLINE=91310689; PubMed=1856229;
RA Wang O., Vandenusen W.J., Petroski C.J., Garaky V.M., Stern A.M.,
RA Friedman P.A.;
RT "Bovine liver aspartyl beta-hydroxylase. Purification and
RT characterization";
RL J. Biol. Chem. 266:14004-14010(1991).
CC -1- FUNCTION: SPECIFICALLY HYDROXYLATES AN ASP OR ASN RESIDUE IN
CC CERTAIN EPIDERMAL GROWTH FACTOR-LIKE (EGF) DOMAINS OF A NUMBER OF
CC PROTEINS.
CC -1- CATALYTIC ACTIVITY: Peptide L-aspartate + 2-oxoglutarate + O(2) =
CC peptide 3-hydroxy-L-aspartate + succinate + CO(2).
CC -1- COFACTOR: IRON.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
CC reticulum.
CC -1- PFM: MIGHT BE PROCESSED TO THE 56 KDA (AA 289-754) OR 52 KDA (AA
CC 311-754) FORMS IN THE LUMEN OF THE ENDOPLASMIC RETICULUM.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M91213; AAA03563.1; -.
CC PIR: A42969; BABOH.
CC InterPro: IPR001440; TPR.
DR Pfam: PF05279; Asp-B-Hydro.N: 1.
DR Pfam: PF05118; Asp.Arg.Hydrox. 1.
KW Oxidoreductase; Dioxigenase; Iron; Transmembrane; Signal-anchor;
KW Endoplasmic reticulum.
FT DOMAIN 1 57
FT TRANSMEM 58 78
FT
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT LUMENAL (POTENTIAL).
FT POLY-GLX.
FT POLY-SER.
FT POLY-LYS.
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 702 702
FT CARBOHYD 702 702
FT SEQUENCE 754 AA: 84998 MW: 369593A1F0B558C8 CCK64;
Query Match 43.1%; Score 43.5; DB 1; Length 754;
Best Local Similarity 55.6%; Pred. No. 42;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Db 82 DYEVLAKANDFRYNLSE 99

```
RESULT 12
ID EM30_ARATH STANDARD: PRT: 1451 AA.
AC 042510: 038883:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pattern formation protein EM30.
GN EM30 OR GNOM OR ATG13980 OR F7L19.7 OR FL6A14.20.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
CC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wassilewskija, and cv. Columbia;
RX MEDLINE=94291195; PubMed=8020095;
RA Shevell D.E., Lee W.-M., Gallmor C.S., Xia G., Feldman K.A.,
RA Chua N.-H.;
RT "EM30 is essential for normal cell division, cell expansion, and
RT cell adhesion in Arabidopsis and encodes a protein that has
RT similarity to Sec7."
RL Cell 77:1051-1062(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia, and cv. CSR-1;
RX MEDLINE=96204508; PubMed=8628228;
RA Busch M., Mayer U., Ueergens G.;
RT "Molecular analysis of the Arabidopsis pattern formation of gene
RT GNOM: gene structure and intergenic complementation."
RL Mol. Gen. Genet. 250:681-691(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA White O., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gutter J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Haller J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.-J., Koo H.L., Kremetskaia I., Kurtz D.B., Kvan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortumí M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
RN [4]
RP -I- FUNCTION: MAY PERFORM A FUNCTION THAT AFFECTS CELL EXPANSION, THE
RP ORIENTATION OF THE PLANE OF CELL DIVISION, THE NUMBER OF CELL
RP DIVISIONS, AND CELL ADHESION THROUGHOUT PLANT DEVELOPMENT.
RN [5]
RP -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
RN [6]
RP -I- TISSUE SPECIFICITY: STEMS, LEAVES, FLOWERS, SILIQUES, FLORAL
RP INFLORESCENCE AND ROOTS.
RN [7]
RP -I- SIMILARITY: Contains 1 Sec7 domain.
RN [8]
RP -----
RN [9]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN [10]
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN [11]
RP the European Bioinformatics Institute. There are no restrictions on its
RN [12]
RP use by non-profit institutions as long as its content is in no way
RN [13]
RP modified and this statement is not removed. Usage by and for commercial
RN [14]
RP entities requires a license agreement (see http://www.isb-sib.ch/announce/
RN [15]
RP or send an email to license@sib-ch.ch).
```

CC EMBL: U56140; AAB01205.1; -
DR EMBL: U56141; AAB01206.1; -
DR EMBL: U36432; AAA91150.1; -
DR EMBL: U36433; AAA91151.1; -
DR EMBL: AC007576; AAD39284.1; -
DR EMBL: AC068197; AAF9403.1; -
DR PIR: S65571; S65571.
DR HSPF: Q99418; IPBV.
DR InterPro: IPR000904; Sec7.
DR Pfam: PF01369; Sec7; 1.
DR SMART: SM00222; Sec7; 1.
DR PROSITE: PS0190; SEC7; 1.
KW Cell adhesion.
FT DOMAIN 537 752 SEC7.
FT MOTIF 658 658 E->K: EMB30-1.
FT CONFLICT 111 111 T->I (IN REF. 2; AAA91150).
FT CONFLICT 867 867 A->G (IN REF. 2; AAA91150).
SQ SEQUENCE 1451 AA; 162618 MW; 666E21C74B426996 CRC64;
Query Match 43.18; Score 43.5; DB 1; Length 1451;
Best Local Similarity 63.28; Pred. No. 83;
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;
DY 1 DLEFIANS-KDFPMSET 18
DB 1302 DLEFIAGSQRDYRMEGT 1320
RESULT 13
UPPS_AQUAE STANDARD; PRT; 231 AA.
AC 067291;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP synthetase)
DE (di-trans-poly-cis-decaprenylcistransferase) (Undecaprenyl diphosphate synthase) (UDS).
DE UPPS OR AQ_1248.
OS Aquifex aeolicus.
GN Aquifex aeolicus.
OC Aquificales; Aquificaceae; Aquifex.
OX NCBI_Taxid=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE=9819666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Jenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.";
RT Nature 392:353-358(1998).
CC -1- FUNCTION: GENERATES UNDECAPRENYL PYROPHOSPHATE (UPP) FROM ISOPENTENYL PYROPHOSPHATE (IPP). UPP IS THE PRECURSOR OF THE CARRIER LIPID FOR PEPTIDOGLYCAN SYNTHESIS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Di-trans-poly-cis-decaprenyl diphosphate + undecaprenyl diphosphate -> diphosphate + di-trans-poly-cis-undecaprenyl diphosphate.
CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: AE000730; AAC07254.1; -
DR PIR: H70407; H70407.
DR InterPro: IPR001441; UPP_synth.
DR Pfam: PF01255; UPP_synthetase; 1.

DR ProDom: PD003461; UPP_synth; 1.
DR TIGRFAMs: TIGR00055; upps; 1.
DR PROSITE: PS01066; UPP_SYNTHETASE; 1.
KW Transferase; Cell division; Cell wall; Peptidoglycan synthesis;
KW Complete proteome.
SQ SEQUENCE 231 AA; 27148 MW; 52B9DE0A4428B0AE CRC64;
Query Match 42.68; Score 43; DB 1; Length 231;
Best Local Similarity 50.08; Pred. No. 15;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
DY 2 LLEFIANSKDFPM 15
DB 112 MEELESDSKDFKNL 125
RESULT 14
40MT_COPUA STANDARD; PRT; 350 AA.
ID 40MT_COPUA
AC Q9LEL5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase (EC 2.1.1.116) (S-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase) (4'-OMT).
DE Coplis japonica (Japanese goldthread).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales; OC Ranunculaceae; Coplis.
OX NCBI_Taxid=3442;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=20390108; PubMed=10811648;
RA Moshige T., Tsujita T., Yamada Y., Sato F.;
RT "Molecular characterization of the S-adenosyl-L-methionine: 3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase involved in licocholine alkaloid biosynthesis in Coplis japonica.";
RL J. Biol. Chem. 275:23398-23405(2000).
CC -1- FUNCTION: CATALYZES THE TRANSFER OF THE METHYL GROUP TO THE 4'-HYDROXYL GROUP OF 3'-HYDROXY-N-METHYLCOCCLAURINE TO FORM RETICULINE.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3'-hydroxy-N-methyl-(S)-coclaurine -> S-adenosyl-L-homocysteine + (S)-reticuline.
CC -1- PATHWAY: CARRIES OUT A STEP IN THE CONVERSION OF 3'-HYDROXY-N-METHYLCOCCLAURINE TO RETICULINE, AN IMPORTANT INTERMEDIATE IN SYNTHESIZING ISOQUINOLINE ALKALOIDS.
CC -1- SUBUNIT: Homodimer.
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL: D29812; BAB08005.1; -
DR InterPro: IPR001601; Methyltransf.
DR InterPro: IPR001077; O-Methyltransf.
DR Pfam: PF00891; Methyltransf_2; 1.
KW Transferase; Methyltransferase.
SQ SEQUENCE 350 AA; 38775 MW; 547835EBCDEF9182 CRC64;
Query Match 42.68; Score 43; DB 1; Length 350;
Best Local Similarity 53.38; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
DY 1 DLEFIANSKDFPM 15
DB 219 DLEFIANSYDLPRNI 233

RESULT 15
 RPOB_CAMJE STANDARD: PRT: 1378 AA.
 ID RPOB_CAMJE 046124; 09P131;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
 beta chain) (RNA polymerase beta subunit).
 GN RPOB OR CJO478.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
 Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 Whitehead S., Barrell B.G.;
 RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
 reveals hypervariable sequences.";
 RT Nature 403:665-668(2000).
 RL [2]
 RN SEQUENCE OF 338-1031 FROM N.A.
 RX MEDLINE=96084944; PubMed=7489896;
 Bustamante V.H., Puente J.L., Sanchez-Lopez F., Bobadilla M.,
 Calva E.;
 RA "Identification of Campylobacter jejuni and C.coli using the rpoB
 gene and a cryptic DNA fragment from C.jejuni.";
 RT Gene 165:1-8(1995).
 RL -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +
 {RNA}(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AL139075; CAB75116.1; -
 DR EMBL: X77304; CA54509.1; -
 DR PIR: A81393; A81393.
 DR PIR: S41868; S41868.
 DR HSSP: Q9KWT7; 1HOM.
 DR InterPro: IPR001572; RNA_pol.B.
 DR Pfam: PF04563; RNA_pol_Rpb2_1; 1.
 DR Pfam: PF04561; RNA_pol_Rpb2_2; 1.
 DR Pfam: PF04565; RNA_pol_Rpb2_3; 1.
 DR Pfam: PF00562; RNA_pol_Rpb2_6; 1.
 DR Pfam: PF04560; RNA_pol_Rpb2_7; 1.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 DR Transfaser: Transcription; DNA-directed RNA polymerase;
 KW Complete proteome.
 KM CONFLICT 338 347 NDIANGVDA -> MTWLMALMP (IN REF. 2).
 FT CONFLICT 558 558 A -> R (IN REF. 2).
 FT CONFLICT 671 671 C -> S (IN REF. 2).
 FT CONFLICT 691 691 A -> R (IN REF. 2).
 SO SEQUENCE 1378 AA; 155915 MW; AB7467C305028EB5 CRC64;

Query Match 41.18; Score 41.5; DB 1; Length 1378;
 Best Local Similarity 47.48; Pred. No. 1.6e+02;
 Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

OY 2 LEEIANSKDFPMSETDL 20
 Db 858 IEEI---TKDIPNVKEDV 873

Search completed: October 1, 2003, 09:57:52
 Job time : 8.58824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEEIAANSKDFPNMSETDL 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mmc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_proteob:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacterioph:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	251	12	Q9ENQ4
2	101	100.0	251	12	Q9ENN2
3	101	100.0	251	12	Q9ENP2
4	101	100.0	251	12	Q9ENP6
5	101	100.0	251	12	Q9ENP1
6	101	100.0	251	12	Q9ENP1
7	101	100.0	251	12	Q9ENP1
8	101	100.0	251	12	Q9ENP1
9	101	100.0	251	12	Q9ENP4
10	101	100.0	251	12	Q9ENP5
11	101	100.0	251	12	Q9ENP7
12	101	100.0	251	12	Q9ENP5
13	101	100.0	251	12	Q9ENP5
14	101	100.0	251	12	Q9ENP6
15	101	100.0	251	12	Q9ENP5
16	101	100.0	251	12	Q9ENN7

17	101	100.0	1124	12	Q84780
18	101	100.0	1161	12	Q05794
19	101	100.0	2216	12	Q9WMA2
20	101	100.0	2218	12	Q67824
21	101	100.0	2225	12	Q9DJ32
22	101	100.0	2227	12	Q9WMA0
23	101	100.0	2227	12	Q9WMA3
24	101	100.0	2227	12	Q67825
25	101	100.0	2227	12	Q9WMA1
26	101	100.0	2227	12	Q67826
27	101	100.0	2227	12	Q8V0N6
28	101	100.0	2227	12	Q9IFN5
29	101	100.0	2227	12	Q9WMA4
30	95	94.1	251	12	Q9ENN9
31	95	94.1	251	12	Q9ENN9
32	95	94.1	251	12	Q9ENN9
33	95	94.1	251	12	Q9ENN9
34	95	94.1	251	12	Q9ENN9
35	95	94.1	251	12	Q9ENN9
36	95	94.1	251	12	Q9ENN9
37	95	94.1	251	12	Q9ENN9
38	95	94.1	251	12	Q9ENN9
39	95	94.1	251	12	Q9ENN9
40	95	94.1	251	12	Q9ENN9
41	95	94.1	251	12	Q9ENN9
42	95	94.1	2218	12	Q67817
43	95	94.1	2227	12	Q8V0V3
44	94	93.1	184	12	Q87092
45	94	93.1	2225	12	Q9DWK1

ALIGNMENTS

RESULT 1

ID Q9ENQ4 PRELIMINARY; PRT; 251 AA.

AC Q9ENQ4;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID-12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A201;

RA Fujiwara K.;

RT "Hepatitis A virus."

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB047659; BAB12167.1; ..

FT NON_TER 1

FT NON_TER 251

SQ SEQUENCE 251 AA; 28720 MW; C3342482862F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20
|||||

DB 95 DLEEIAANSKDFPNMSETDL 114

RESULT 2

ID Q9ENN2 PRELIMINARY; PRT; 251 AA.

AC Q9ENN2;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polypeptide (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A9;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047681; BAB12189.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C334248282E19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 3

Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC O9ENP2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A5;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047671; BAB12179.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 4

Q9ENP6 PRELIMINARY; PRT; 251 AA.
AC O9ENP6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A162;
RA Fujiwara K.;

RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047657; BAB12165.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 5

Q9ENR1 PRELIMINARY; PRT; 251 AA.
AC O9ENR1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A1;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047652; BAB12160.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 6

Q9ENP1 PRELIMINARY; PRT; 251 AA.
AC O9ENP1;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A503;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047672; BAB12180.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
|||||
95 DLEETIANSKDFPNMSETDL 114

Db

RESULT 7

O9ENQ1 PRELIMINARY; PRT; 251 AA.

AC O9ENQ1: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
RA Fujikura K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047662; BAB12170.1; -.
FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
|||||
95 DLEETIANSKDFPNMSETDL 114

Db

RESULT 8

O9ENQ9 PRELIMINARY; PRT; 251 AA.

AC O9ENQ9: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
RA Fujikura K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
|||||
95 DLEETIANSKDFPNMSETDL 114

Db

RESULT 9

O9ENNA

ID O9ENN4 PRELIMINARY; PRT; 251 AA.

AC O9ENN4: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
RA Fujikura K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
|||||
95 DLEETIANSKDFPNMSETDL 114

Db

RESULT 10

O9ENP5 PRELIMINARY; PRT; 251 AA.

AC O9ENP5: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
RA Fujikura K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
|||||
95 DLEETIANSKDFPNMSETDL 114

Db

RESULT 11

O9ENP7 PRELIMINARY; PRT; 251 AA.

AC O9ENP7: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A306;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||

Db 95 DLEEIAANSKDFPMSETDL 114

SOUT 12

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A20;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||

Db 95 DLEEIAANSKDFPMSETDL 114

RESULT 13

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A303;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047664; BAB12172.1; -.
FT NON_TER 1
FT NON_TER 1

FT NON_TER 251
SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CASC1A CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||

Db 95 DLEEIAANSKDFPMSETDL 114

RESULT 14

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A713;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||

Db 95 DLEEIAANSKDFPMSETDL 114

RESULT 15

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A75;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047678; BAB12186.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28658 MW; 98E8BEDD00B2EDF10 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||

Thu Oct 2 09:14:00 2003

us-09-171-432a-45.rspt

Page 5

Db 95 DLEBIANSKDPNNSETDL 114

Search completed: October 1, 2003, 10:02:40
Job time : 33.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 : Search time 12.7059 Seconds
(without alignments)
66,600 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEETANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	3	US-08-475-886-2
2	101	100.0	2227	3	US-08-475-886-4
3	101	100.0	2227	3	US-08-475-886-6
4	101	100.0	2227	3	US-08-397-232-2
5	101	100.0	2227	3	US-08-397-232-4
6	101	100.0	2227	3	US-09-171-387-2
7	101	100.0	2227	3	US-09-653-499-2
8	101	100.0	2227	4	US-09-653-499-4
9	101	100.0	2227	4	US-09-653-499-6
10	44	43.6	907	3	US-08-938-830-26
11	44	43.6	907	3	US-09-020-222-26
12	43	42.6	1201	3	US-09-098-901-2
13	42	41.6	365	2	US-08-983-352-2
14	42	41.6	416	4	US-09-107-532A-5551
15	41	40.6	376	3	US-09-200-965-2
16	41	40.6	508	4	US-09-252-991A-18910
17	41	40.6	528	4	US-09-356-806-8
18	40.5	40.1	191	4	US-09-198-452A-1095
19	40	39.6	496	4	US-09-339-159B-281
20	39.5	39.1	224	4	US-09-134-001C-3648
21	39	38.6	50	1	US-08-127-351-14
22	39	38.6	50	1	US-08-480-367B-14
23	39	38.6	50	1	US-08-487-221A-14
24	39	38.6	50	1	US-08-480-370-14
25	39	38.6	135	4	US-09-198-452A-544
26	39	38.6	168	4	US-09-134-001C-4664
27	39	38.6	288	4	US-09-107-532A-6783

28	39	38.6	352	3	US-09-286-691-26	Sequence 26, Appl
29	39	38.6	352	3	US-09-687-147-26	Sequence 26, Appl
30	39	38.6	387	4	US-09-314-847A-4	Sequence 4, Appl
31	39	38.6	387	4	US-09-570-778A-2	Sequence 2, Appl
32	39	38.6	387	4	US-09-570-778A-7	Sequence 7, Appl
33	39	38.6	387	4	US-09-570-778A-8	Sequence 8, Appl
34	39	38.6	387	4	US-09-570-778A-9	Sequence 9, Appl
35	39	38.6	387	4	US-09-991-138-2	Sequence 2, Appl
36	39	38.6	387	4	US-09-991-138-7	Sequence 7, Appl
37	39	38.6	387	4	US-09-991-138-8	Sequence 8, Appl
38	39	38.6	387	4	US-09-991-138-9	Sequence 9, Appl
39	39	38.6	391	4	US-09-482-273-151	Sequence 151, App
40	39	38.6	396	1	US-08-430-024-2	Sequence 2, Appl
41	39	38.6	396	1	US-08-782-009-2	Sequence 2, Appl
42	39	38.6	396	3	US-09-017-302-2	Sequence 2, Appl
43	39	38.6	414	4	US-09-482-273-239	Sequence 239, App
44	39	38.6	984	1	US-08-242-932-2	Sequence 2, Appl
45	39	38.6	984	1	US-08-714-481-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-08-475-886-2
Sequence 2, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY      1 DLEETANSKDFPNMSETDL 20
DB      931 DLEETANSKDFPNMSETDL 950
RESULT 2
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
```

```
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

RESULT 3

```
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

RESULT 4

```
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
```

```
; TYPE: PRF
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

RESULT 5

```
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

RESULT 6

```
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
```

```
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; APPLICANT: EMERSON, SUZANNE, U.;
; APPLICANT: PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
```


PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-42290S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262052
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 8
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U

APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262052
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 9
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262052
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20
DB 931 DLEETIANSKDFPNMSETDL 950

RESULT 10
US-08-938-830-26
Sequence 26, Application US/08938830
Patent No. 6040437
GENERAL INFORMATION:
APPLICANT: Lasky, Laurence A.
APPLICANT: Dowbenko, Donald J.
TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage

```

? TITLE OF INVENTION: Furrow-Associated Proteins (PSTrips)
? NUMBER OF SEQUENCES: 73
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/938,830
? FILING DATE:
? CLASSIFICATION: 530
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/798419
? FILING DATE: 07-FEB-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Dreger, Ginger R.
? REGISTRATION NUMBER: 33,055
? REFERENCE/DOCKET NUMBER: P1066P1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-3216
? TELEFAX: 650/952-9881
? INFORMATION FOR SEQ ID NO: 26:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 907 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
US-08-938-830-26

Query Match      43.6%; Score 44; DB 3; Length 907;
Best Local Similarity 47.1%; Pred. No. 45;
Matches      8; Conservative 4; Mismatches      5; Indels      0; Gaps      0;

QY      2 LELTANSKDPNMSSET 18
       |::|::|::|::|
Db      49 LQELASSADPIEYGST 65

RESULT 11
US-09-020-222-26
? Sequence 26, Application US/09020222
? Patent No. 6111073
? GENERAL INFORMATION:
? APPLICANT: Laskey, Laurence A.
? TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage
? TITLE OF INVENTION: Furrow-Associated Proteins (PSTrips)
? NUMBER OF SEQUENCES: 27
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: WinPatIn (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/020,222
? FILING DATE: 06-Feb-1998
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/798419
? FILING DATE: 02/07/1997
? ATTORNEY/AGENT INFORMATION:

```

```

NAME: Dreger, Glander R.
REGISTRATION NUMBER: 33,055
REFERENCE/DOCKET NUMBER: P1066r1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-3216
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-020-222-26

Query Match      43.6%: Score 44; DB 3; Length 907;
Best Local Similarity 47.1%: Pred. NO. 45;
Matches      8; Conservative      4; Mismatches      5; Indels      0; Gaps      0;

QY      2 LEEIANSKDFPNMSETD 18
Db      49 LQELASSADIPYVGST 65

RESULT 12
US-09-098-901-2
Sequence 2, Application US/09098901B
Patent No. 6218144
GENERAL INFORMATION:
APPLICANT: Scott, Matthew
APPLICANT: Sisson, John C.
TITLE OF INVENTION: Costal2 Genes and their Uses
FILE REFERENCE: SUN-65P
CURRENT APPLICATION NUMBER: US/09/098,901B
CURRENT FILING DATE: 1998-06-17
EARLIER APPLICATION NUMBER: 60/051,347
EARLIER FILING DATE: 1997-06-30
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 1201
TYPE: PRT
ORGANISM: D. Melanogaster
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1201)
OTHER INFORMATION: xaa = Any Amino Acid
US-09-098-901-2

Query Match      42.6%: Score 43; DB 3; Length 1201;
Best Local Similarity 50.0%: Pred. NO. 94;
Matches      8; Conservative      4; Mismatches      4; Indels      0; Gaps      0;

QY      4 EIAANSKDFPNMSETD 19
Db      48 QVPAAEKDNPSISSETD 63

RESULT 13
US-08-983-352-2
Sequence 2, Application US/08983352
Patent No. 5928859
GENERAL INFORMATION:
APPLICANT: Nicklin, Stephen
APPLICANT: Blinks, Peter R.
APPLICANT: Bruce, Neil C.
APPLICANT: French, Christopher E.
TITLE OF INVENTION: DETECTION AND BIODEGRADATION OF
TITLE OF INVENTION: EXPLOSIVES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NIXON & VANDERHAYE P.C.
STREET: 1100 NO. 5928859th Glebe Rd. 8th floor
CITY: Arlington
STATE: VA

```

COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,352
FILING DATE: 02-FEB-1998
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/01629
FILING DATE: 08-JUL-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9514138.8
FILING DATE: 11-JUL-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9607220.2
FILING DATE: 04-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 124-617
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-983-352-2

Query Match 41.6%; Score 42; DB 2; Length 365;
Best Local Similarity 47.4%; Pred. No. 33;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 LEEIANSKDPNMSSTD 20
DB 258 IEELAKRGIAVLMSTDL 276

RESULT 14
US-09-107-532A-5551
Sequence 5551, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucelte-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENEOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: CTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5551:
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (b) LOCATION 1...416
SEQUENCE DESCRIPTION: SEQ ID NO: 5551:
US-09-107-532A-5551

Query Match 41.6%; Score 42; DB 4; Length 416;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 LEEIANSKDPNPM 15
DB 117 ERYLTNTKDPNL 129

RESULT 15
US-09-200-965-2
Sequence 2, Application US/09200965
Patent No. 6134422
GENERAL INFORMATION:
APPLICANT: ROSEN, ET AL.
TITLE OF INVENTION: Thrombin Inhibitor
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/200,965
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/171,817
FILING DATE: 22-DECEMBER-1993
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-47
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN

Thu Oct 2 09:13:59 2003

us-09-171-432a-45.raii

Page 6

US-09-200-965-2

Query Match	40.6%	Score 41	DB 3	Length 376
Best Local Similarity	33.3%	Pred. NO	50	
Matches 10	Conservative 3	Mismatches 7	Indels 10	Gaps 1

```

QY      1 DLEEIAAN-----SKDFPNSETDL 20
        1:1: | | | | |
DB      281 DMESVLRNIGMTDAFELGKADFSGNSQTDL 310

```

Search completed: October 1, 2003, 10:06:32
Job time : 13.7059 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds

(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEIRIANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubpaa/PC1_NEM_PUB.pep:*

3: /cgn2_6/ptodata/1/pubpaa/US06_NEM_PUB.pep:*

4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubpaa/US07_NEM_PUB.pep:*

6: /cgn2_6/ptodata/1/pubpaa/PC1_NEM_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubpaa/US08_NEM_PUB.pep:*

8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubpaa/US09_NEM_PUB.pep:*

13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubpaa/US10_NEM_PUB.pep:*

17: /cgn2_6/ptodata/1/pubpaa/US60_NEM_PUB.pep:*

18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	US-10-272-459-41	Sequence 41, Appl
3	101	100.0	2227	US-09-929-955-12	Sequence 12, Appl
4	101	100.0	2227	US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	US-10-135-988-4	Sequence 6, Appl
7	101	100.0	2227	US-10-135-988-6	Sequence 27, Appl
8	44	43.6	401	US-10-195-144-27	Sequence 34, Appl
9	42	41.6	538	US-09-769-744A-34	Sequence 41631, A
10	41	40.6	72	US-09-864-761-41831	Sequence 1440, Ap
11	41	40.6	459	US-09-925-300-1440	Sequence 8, Appl
12	40	39.6	528	US-10-205-522-8	Sequence 1, Appl
13	40	39.6	284	US-09-810-997-1	Sequence 1, Appl
14	40	39.6	284	US-10-174-209-1	Sequence 11280, A
15	40	39.6	381	US-09-815-242-11280	

16	40	39.6	499	9	US-09-864-761-35385	Sequence 35385, A
17	40	39.6	569	15	US-10-156-761-12273	Sequence 12273, A
18	40	39.6	715	12	US-10-032-585-7002	Sequence 7002, Ap
19	40	39.6	786	12	US-10-263-568-8	Sequence 8, Appl
20	40	39.6	1230	12	US-10-205-219-44	Sequence 44, Appl
21	40	39.6	1230	12	US-10-205-219-97	Sequence 97, Appl
22	39	38.6	337	12	US-10-174-209-6	Sequence 6, Appl
23	39	38.6	331	12	US-09-803-286A-10	Sequence 10, Appl
24	39	38.6	341	9	US-10-166-225A-87	Sequence 87, Appl
25	39	38.6	341	9	US-09-815-242-10657	Sequence 10657, A
26	39	38.6	387	11	US-09-991-138-2	Sequence 2, Appl
27	39	38.6	387	11	US-09-991-138-7	Sequence 7, Appl
28	39	38.6	387	11	US-09-991-138-8	Sequence 8, Appl
29	39	38.6	387	11	US-09-991-138-9	Sequence 9, Appl
30	39	38.6	387	11	US-10-037-677-4	Sequence 4, Appl
31	39	38.6	391	11	US-09-984-271-151	Sequence 151, App
32	39	38.6	396	9	US-09-815-242-10387	Sequence 10387, A
33	39	38.6	414	11	US-09-984-271-239	Sequence 239, App
34	39	38.6	417	9	US-09-815-242-10492	Sequence 10492, A
35	39	38.6	420	10	US-09-764-868-912	Sequence 912, App
36	39	38.6	569	15	US-10-034-934-87	Sequence 87, Appl
37	39	38.6	722	15	US-10-222-668-4	Sequence 4, Appl
38	39	38.6	978	10	US-09-893-817-4	Sequence 4, Appl
39	39	38.6	978	10	US-09-893-817-8	Sequence 8, Appl
40	39	38.6	1111	9	US-09-815-242-12955	Sequence 12955, A
41	39	38.6	1247	9	US-09-803-286A-2	Sequence 2, Appl
42	38.5	38.1	383	12	US-10-037-585-7513	Sequence 7513, App
43	38	37.6	89	12	US-10-440-066-36	Sequence 36, Appl
44	38	37.6	89	12	US-10-440-066-37	Sequence 37, Appl
45	38	37.6	89	12	US-10-440-066-38	Sequence 38, Appl

ALIGNMENTS

RESULT 1

US-10-272-459-45

; Sequence 45, Application US/10272459

; Publication No. US20030124517A1

; GENERAL INFORMATION:

; APPLICANT: PICHUANES, Sergio

; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

; TITLE OF INVENTION: PROTEINS AND USES THEREOF

; FILE REFERENCE: P17955.002 / 2301-17955

; CURRENT APPLICATION NUMBER: US/10/272,459

; CURRENT FILING DATE: 2002-10-15

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 45

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein

; OTHER INFORMATION: of 38.8 kDa

US-10-272-459-45

Query Match 100.0%; Score 101; DB 15; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEIRIANSKDFPNMSETDL 20

|||||

DB 303 DLEIRIANSKDFPNMSETDL 322

RESULT 2

US-10-272-459-41

; Sequence 41, Application US/10272459

; Publication No. US20030124517A1

; GENERAL INFORMATION:

; APPLICANT: PICHUANES, Sergio

; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```

; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; US-10-272-459-41
```

```

Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEEIANSKDFPMSETDL 20
        |||
Db       931 DLEEIANSKDFPMSETDL 950
```

```

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
; 09-929-955-12
```

```

Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEEIANSKDFPMSETDL 20
        |||
Db       931 DLEEIANSKDFPMSETDL 950
```

```

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
; US-10-104-966-12
```

```

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEEIANSKDFPMSETDL 20
        |||
Db       931 DLEEIANSKDFPMSETDL 950
```

```

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HN-175
; US-10-135-988-2
```

```

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEEIANSKDFPMSETDL 20
        |||
Db       931 DLEEIANSKDFPMSETDL 950
```

```

RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRF
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 14;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNMSETDL 20
DB 931 DLEETANSKDFPNMSETDL 950

RESULT 7
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176669A1

;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262053
;; CURRENT APPLICATION NUMBER: US/10/135,988
;; CURRENT FILING DATE: 2002-04-29
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRF
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 14;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNMSETDL 20
DB 931 DLEETANSKDFPNMSETDL 950

RESULT 8
US-10-195-144-27
Sequence 27, Application US/10195144
Publication No. US20030126646A1
;; GENERAL INFORMATION:
;; APPLICANT: BROWN, GREGORY G.
;; APPLICANT: FORMANOVA, NATASHA
;; APPLICANT: DENDY, CHARLES
;; APPLICANT: LANDRY, BENOIT S.
;; APPLICANT: CHEUNG, WING
;; APPLICANT: JIN, HUA
;; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
;; FILE REFERENCE: 16313-0136
;; CURRENT APPLICATION NUMBER: US/10/195,144
;; CURRENT FILING DATE: 2002-10-01
;; PRIOR APPLICATION NUMBER: 60/305,026
;; PRIOR FILING DATE: 2001-07-12
;; PRIOR APPLICATION NUMBER: 60/305,363
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: 60/308,736

;; PRIOR FILING DATE: 2001-07-30
;; NUMBER OF SEQ ID NOS: 128
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 27
;; LENGTH: 401
;; TYPE: PRF
;; ORGANISM: Raphanus sativum
US-10-195-144-27

Query Match 43.6%; Score 44; DB 15; Length 401;
Best Local Similarity 56.2%; Pred. No. 44;
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNM 16
DB 9 ELDEETALSKEFMSLS 24

RESULT 9
US-09-769-744A-34
Sequence 34, Application US/09769744A
Publication No. US20030134407A1

;; GENERAL INFORMATION:
;; APPLICANT: Le Page, Richard WF
;; APPLICANT: Wells, Jeremy M
;; APPLICANT: Hanniffy, Sean B
;; APPLICANT: Hansbro, Philip M
;; TITLE OF INVENTION: Proteins
;; FILE REFERENCE: PWC/P21122WO
;; CURRENT APPLICATION NUMBER: US/09/769,744A
;; CURRENT FILING DATE: 2001-01-26
;; PRIOR APPLICATION NUMBER: PCT/GB99/02452
;; PRIOR FILING DATE: 1999-07-27
;; PRIOR APPLICATION NUMBER: GB 9816336.3
;; PRIOR FILING DATE: 1998-07-27
;; PRIOR APPLICATION NUMBER: US 60/125329
;; PRIOR FILING DATE: 1999-03-19
;; NUMBER OF SEQ ID NOS: 196
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 34
;; LENGTH: 538
;; TYPE: PRF
;; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-34

Query Match 41.6%; Score 42; DB 12; Length 538;
Best Local Similarity 53.8%; Pred. No. 1.3e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 LEEETANSKDFPN 14
DB 436 LEOVGSLVKDFPN 448

RESULT 10
US-09-864-761-41831
Sequence 41831, Application US/09864761
Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecm1ca-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366

;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 41831
;; LENGTH: 72
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC025488.1
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
;; OTHER INFORMATION: EST_HUMAN HIT: BF689099.1, EVALUE 7.00e-34
;; OTHER INFORMATION: SWISSPROT HIT: P06133, EVALUE 1.00e-37
;; -09-864-761-41831

Query Match 40.6%; Score 41; DB 9; Length 72;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 AANSKDFPMSETDL 20
DB 47 AAVSLDFHTMSSTD 61

RESULT 11
US-09-925-300-1440
;; Sequence 1440, Application US/09925300
;; Patent No. US20020151681A1
;; GENERAL INFORMATION:
;; APPLICANT: Craig Rosen,
;; APPLICANT: Steve Ruben
;; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;; FILE REFERENCE: PA101
;; PRIOR APPLICATION NUMBER: US/09/925,300
;; CURRENT FILING DATE: 2001-08-10
;; PRIOR APPLICATION NUMBER: PCT/US00/05988
;; PRIOR FILING DATE: 2000-03-08
;; PRIOR APPLICATION NUMBER: 60/124,270

;; PRIOR FILING DATE: 1999-03-12
;; NUMBER OF SEQ ID NOS: 1890
;; SOFTWARE: Patent In Ver. 2.0
;; SEQ ID NO 1440
;; LENGTH: 459
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-925-300-1440

Query Match 40.6%; Score 41; DB 10; Length 459;
Best Local Similarity 33.3%; Pred. No. 1.6e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 10; Gaps 1;

OY 1 DLEBIAAN-----SKDFPMSETDL 20
DB 364 DMESVLRNLGWTDAFELGKADFGSMQSTD 393

RESULT 12
US-10-205-522-8
;; Sequence 8, Application US/10205522
;; Publication No. US2003007629A1
;; GENERAL INFORMATION:
;; APPLICANT: Penny, Laura
;; APPLICANT: Galvin, Margaret
;; APPLICANT: Miller, Andrew
;; APPLICANT: Reidy, Michael
;; TITLE OF INVENTION: Genotyping Human
;; TITLE OF INVENTION: UDP-Glucuronosyltransferase 2B4 (UGT2B4), 2B7 (UGT2B7) and
;; FILE REFERENCE: SEQ-22PRV2
;; CURRENT APPLICATION NUMBER: US/10/205,522
;; CURRENT FILING DATE: 2002-07-24
;; PRIOR APPLICATION NUMBER: US/09/356,806
;; PRIOR FILING DATE: 1999-07-20
;; NUMBER OF SEQ ID NOS: 164
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 8
;; LENGTH: 528
;; TYPE: PRT
;; ORGANISM: H. sapiens
US-10-205-522-8

Query Match 40.6%; Score 41; DB 15; Length 528;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 AANSKDFPMSETDL 20
DB 411 AAVSLDFHTMSSTD 425

RESULT 13
US-09-810-997-1
;; Sequence 1, Application US/09810997
;; Patent No. US20020007501A1
;; GENERAL INFORMATION:
;; APPLICANT: Song, Xiaoling
;; APPLICANT: Fan, Hao
;; APPLICANT: Wei, Zhong-Min
;; TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
;; FILE REFERENCE: 21829/62
;; CURRENT APPLICATION NUMBER: US/09/810,997
;; CURRENT FILING DATE: 2001-03-16
;; PRIOR APPLICATION NUMBER: 60/191,649
;; PRIOR FILING DATE: 2000-03-23
;; PRIOR APPLICATION NUMBER: 60/250,710
;; PRIOR FILING DATE: 2000-12-01
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: Patent In Ver. 2.1
;; SEQ ID NO 1
;; LENGTH: 284

TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-810-997-1

Query Match
Best Local Similarity 39.6%; Score 40; DB 9; Length 284;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 EEIANSKDFPNMSETDL 20
DB 174 GRIDVFSKDFDNIAEVEL 191

RESULT 14
US-10-174-209-1
Sequence 1, Application US/10174209
Publication No. US2003017526A1

GENERAL INFORMATION:

APPLICANT: Song, Xiaoling

APPLICANT: Bartola, Pauline A.

APPLICANT: Linderroth, No. US20030177526A1 A.

APPLICANT: Fan, Hao

APPLICANT: Wei, Zhong-Min

TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: 21829/211

CURRENT APPLICATION NUMBER: US/10/174,209

CURRENT FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: 60/335,776

PRIOR FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: 09/810,997

PRIOR FILING DATE: 2001-03-16

NUMBER OF SEQ ID NOS: 86

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 284

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-174-209-1

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 284;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 EEIANSKDFPNMSETDL 20
DB 174 GRIDVFSKDFDNIAEVEL 191

RESULT 15
US-09-815-242-11280

Sequence 11280, Application US/09815242
Patent No. US20020061569A1

GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Treweek, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11280
LENGTH: 381
TYPE: PRT
ORGANISM: Haemophilus influenzae
US-09-815-242-11280

Query Match
Best Local Similarity 39.6%; Score 40; DB 9; Length 381;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNMSETD 19
DB 41 DLEIANSKDFPNMSETD 59

Search completed: October 1, 2003, 10:37:52
Job time : 24.1176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq.19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	AAW42929	Immunogenic Hepat
2	96	100.0	20	AAW42929	Synthetic HAV P2A
3	96	100.0	25	AAW42969	Immunogenic Hepat
4	96	100.0	25	AAW42969	Synthetic HAV P2A
5	96	100.0	22	AAW42969	Sequence of viral
6	96	100.0	22	AAW42969	Attenuated hepati
7	96	100.0	22	AAW42969	Hepatitis A virus
8	96	100.0	22	AAW42969	Amino acid sequenc
9	96	100.0	22	AAW42969	Amino acid sequenc

10	96	100.0	22	22	21	AAW42929
11	96	100.0	22	22	23	AAW42929
12	96	100.0	22	22	23	AAW42929
13	96	100.0	22	22	23	AAW42929
14	96	100.0	22	22	23	AAW42929
15	96	100.0	22	22	23	AAW42929
16	96	100.0	22	22	23	AAW42929
17	96	100.0	22	22	23	AAW42929
18	96	100.0	22	22	23	AAW42929
19	96	100.0	22	22	23	AAW42929
20	96	100.0	22	22	23	AAW42929
21	96	100.0	22	22	23	AAW42929
22	96	100.0	22	22	23	AAW42929
23	96	100.0	22	22	23	AAW42929
24	96	100.0	22	22	23	AAW42929
25	96	100.0	22	22	23	AAW42929
26	96	100.0	22	22	23	AAW42929
27	96	100.0	22	22	23	AAW42929
28	96	100.0	22	22	23	AAW42929
29	96	100.0	22	22	23	AAW42929
30	96	100.0	22	22	23	AAW42929
31	96	100.0	22	22	23	AAW42929
32	96	100.0	22	22	23	AAW42929
33	96	100.0	22	22	23	AAW42929
34	96	100.0	22	22	23	AAW42929
35	96	100.0	22	22	23	AAW42929
36	96	100.0	22	22	23	AAW42929
37	96	100.0	22	22	23	AAW42929
38	96	100.0	22	22	23	AAW42929
39	96	100.0	22	22	23	AAW42929
40	96	100.0	22	22	23	AAW42929
41	96	100.0	22	22	23	AAW42929
42	96	100.0	22	22	23	AAW42929
43	96	100.0	22	22	23	AAW42929
44	96	100.0	22	22	23	AAW42929
45	96	100.0	22	22	23	AAW42929

ALIGNMENTS

RESULT 1
AAW42929 standard; peptide; 20 AA.
AAW42929:
28-APR-1998 (first entry)
Immunogenic Hepatitis A virus peptide YK-1331.
Immunogenic peptide; immunogenic epitope; P2A protein;
Immune response; antibody.
Synthetic.
Hepatitis A virus.
WO9740147-A1.
30-OCT-1997.
18-APR-1997; 97WO-US06891.
19-APR-1996; 96US-0015644.
(US) US DEPT HEALTH & HUMAN SERVICES.
Fields HA, Khudjakov YE;
WPI; 1997-535831/49.
Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an
Immune response to HAV in a mammal or to detect the presence of

Amino acid sequenc
Wild-type Hepatiti
Hepatitis A virus
Attenuated Hepatit
Hepatitis A virus
Wild type human he
Attenuated (pass3
Attenuated hepati
Propionibacterium
Streptococcus poly
Shrimp white spot
Streptococcus poly
Lactococcus lactis
Propionibacterium
Saccharomyces cere
Amino acid sequenc
Human kinase (PKIN
Human steroid and
Human steroid and
Drosophila melanog
Novel human diagno
Protein which is s
Mouse CD14 protein
A murine CD14 prot
Pseudomonas aerugi
Chlamydia trachoma
Amino acid sequenc
Human lung cancer
Human clone L761P
Novel human diagno
Cyanobacterial phy
Novel human diagno
C glutamicum prote
Propionibacterium
Propionibacterium
Human liver peptid

PT antibodies against HAV in a mammal
XX
XX Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 961-980, and has a reactivity of 27.1% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal.
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 96; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KINLADRLGLSGVOEIKQ 20
1 KINLADRLGLSGVOEIKQ 20
Db
RESULT 2
AAB69446
ID AAB69446 standard; Peptide: 20 AA.
XX
AC AAB69446;
XX
XX 20-APR-2001 (first entry)
DT
XX
XX Synthetic HAV P2A peptide, SEQ ID NO: 46.
DE
XX
KM Hepatitis A virus; HAV, immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
XX 15-JUL-1999; 99US-0144412.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Fields HA, Khudyakov YE;
XX
XX WPI; 2001-112681/12.
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
XX
PS Claim 13; Page 97; 130pp; English.
XX
XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC Igm antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the Igm antibody reactivity.
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 96; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KINLADRLGLSGVOEIKQ 20
1 KINLADRLGLSGVOEIKQ 20
Db
RESULT 3
AAW42969
ID AAW42969 standard; peptide: 25 AA.
XX
AC AAW42969;
XX
XX 28-APR-1998 (first entry)
DT
XX
XX Immunogenic Hepatitis A virus peptide YK-1757.
DE
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein;
KW immune response; antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
XX WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US06891.
XX
XX 19-APR-1996; 96US-0015644.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Fields HA, Khudyakov YE;
XX
XX WPI; 1997-535831/49.
XX
XX
PT Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an
PT immune response to HAV in a mammal or to detect the presence of
PT antibodies against HAV in a mammal
XX
XX
PS Claim 18; Page 112; 140pp; English.
XX
XX The present immunogenic peptide corresponds to an immunogenic
CC epitope of the Hepatitis A virus (HAV). The peptide is substantially
CC similar to a portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. Compositions containing the
CC peptide can be used to induce an immune response to HAV in a mammal.
CC The peptide can also be used to detect the presence of antibodies
CC against HAV in mammalian serum. The peptide can also be used to make an
CC antibody against HAV by administering the peptide to a mammal.
XX
SQ Sequence 25 AA;
Query Match 100.0%; Score 96; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KINLADRLGLSGVOEIKQ 20
6 KINLADRLGLSGVOEIKQ 25
Db
RESULT 4

AA69448	standard; Peptide; 25 AA.
AA69448	
20-APR-2001	(first entry)
Synthetic HAV P2A peptide, SEQ ID NO: 48.	
Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine; antigen; major structural capsid polypeptide; HAV antibody detection.	
Hepatitis A virus.	
Synthetic.	
MO200105824-A2.	
25-JAN-2001.	
14-JUL-2000; 2000WO-US19267.	
15-JUL-1999; 99US-0144412.	
(USSH) US DEPT HEALTH & HUMAN SERVICES.	
Fields HA, Khudyakov YE;	
WPI; 2001-113681/12.	
Synthetic peptides used as antigen sources for enzyme immunoassays detecting anti-hepatitis A virus and as vaccines -	
Claim 13; Page 99; 130pp; English.	
The present sequence is one of a number of synthetic peptides which are immunoreactive with hepatitis A virus (HAV) antibodies. The peptides comprise antigenic epitopes of the major structural capsid polypeptides or non-structural polypeptides of HAV with one or more glutamine molecules at the carboxy end of the peptide. The peptides are used to detect the presence of antibodies against HAV in mammalian serum, to detect the presence of HAV in a human or animal through the binding of the peptide to an antibody, to detect acute phase infection by detecting Igm antibodies in mammalian serum and detecting convalescence in a mammal. The peptides are used to detect or quantify HAV antibodies in samples in clinical or research-based assays using immunoblotting, fluorescence in situ hybridization analysis, gel-mobility shift assays, tracking of radioactive or bioluminescent markers, chromatography or electrophoresis. The peptides are used to induce an immune response to HAV when administered to a human or animal. Glutamine at the carboxy end of the peptides enhances the Igm antibody reactivity.	
Sequence 25 AA:	
Query Match	100.0%; Score 96; DB 22; Length 25;
Best Local Similarity	100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
1 KINTADRLGLSGVOEIKQ 20	
6 KINTADRLGLSGVOEIKQ 25	
RESULT 5	
AA60066	
AA60066 standard; Protein; 2227 AA.	
AA60066;	
25-MAR-2003 (updated)	
26-JUN-1991 (first entry)	
Sequence of viral I434 polypeptide encoded by the complete nucleotide sequence of the HAV genome.	

XX		Diagnosis; vaccine; passive immunotherapy.
XX		Hepatitis A virus.
OS		
XX	Key	Location/Qualifiers
PH	Region	1..245
FT	Region	/label- p1.1A
FT	Region	246..491
FT	Region	/label- 1B
FT	Region	492..836
FT	Region	/label- 1C
FT	Region	837..980
FT	Region	/label- p2.2A
FT	Region	981..1076
FT	Region	/label- 2B
FT	Region	1077..1422
FT	Region	/label- 2C
FT	Region	1423..1484
FT	Region	/label- p3.3A
FT	Region	1485..1507
FT	Region	/label- 3B
FT	Region	1508..1678
FT	Region	/label- 3C
FT	Region	1679..2227
FT	Region	/label- 3D
XX		
PN	EP199480-A.	
PD	29-OCT-1986.	
PE	03-APR-1986;	86EP-0302465.
PR	03-APR-1985;	85US-0719329.
PA	(CHIR) CHIRON CORP.	
P1	Dina D, Potter SJ, Vannest GA, Caput D;	
DR	WPI; 1986-286213/44.	
DR	N-PSDB; AAN60080.	
XX		
PT	Hepatitis A virus nucleotide sequence and polypeptide - and use	
PT	In prodn. of vaccines and diagnostic probes	
XX		
PS	Claim 5; Fig 1; 18pp; English.	
CC	AAN60080 and oligonucleotide fragments are useful in detection of	
CC	hepatitis A virus; transformed hosts may be used for expression of	
CC	polypeptides and fragments useful in vaccines without risk of	
CC	infection by the virus or in prodn. of particles which are capable	
CC	of inducing immunocompetent B cells for passive immunotherapy. Pref.	
CC	epitope is derived from AAs 445-657 or 792-848 of the HAV	
CC	polypeptide sequence (AAP60066).	
CC	(Updated on 25-MAR-2003 to correct PA field.)	
XX		
SQ	Sequence	2227 AA;
OY	Query Match	100.0%; Score 96; DB 7; Length 2227;
	Best Local Similarity	100.0%; Pred. No. 2.2e-07;
	Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 KINLADRLGLSGVOEIKQ 20	
	961 KINLADRLGLSGVOEIKQ 980	
RESULT 6		
AA005697		
ID	AA005697 standard; protein: 2227 AA.	
AC	AA005697;	
XX		

```

DE 25-MAR-2003 (updated)
DF 15-AUG-1990 (first entry)
XX
XX Attenuated hepatitis A virus.
DE
XX
XX Hepatitis A virus; vaccine; attenuated.
KM
XX
OS Hepatitis A virus, strain HM-175.
XX
XX
XX Key Location/Qualifiers
FH 1..23
FT Region
FT /label-VP4 = 1A
FT 24..245
FT /label-VP2 = 1B
FT 246..491
FT /label-VP3 = 1C
FT 492..791
FT Region
FT /label-VP1 = 1D
FT 792..980
FT /label-2A
FT 981..1087
FT /label-2B
FT 1088..1422
FT Region
FT /label-2C
FT 1423..1496
FT /label-3A
FT 1497..1519
FT /label-3B = VPg
FT 1520..1738
FT Region
FT /label-3C
FT 1739..2227
FT /label-3D
FT
XX
XX US4894228-A.
PN
XX
XX 16-JAN-1990.
PD
XX
XX 12-JUL-1988; 88US-0217824.
PE
XX
XX 12-JUL-1988; 88US-0217824.
XX
XX 12-JUL-1988; 88US-0217824.
PR 12-JUL-1988; 88US-0652967.
XX
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
PA
XX
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM,
PI Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
XX
XX N-PSDB; AMQ03512.
XX
XX
XX Vaccine against hepatitis A virus infection - comprises novel
XX attenuated hepatitis A virus strain.
XX
XX
XX Claim 1; Fig 1; 18pp; English.
XX
XX
XX The attenuated HAV is useful for inducing protective immunity against
XX HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
XX several nucleotide changes distributed throughout the genome, is
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is
XX suitable for use as an HAV vaccine. It is noted that not all the changes
XX are necessary for attenuation and use as a vaccine.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX
XX Sequence 2227 AA;
SQ

```

RESULT	7
AAM34074	ID AAM34074 standard; Protein: 2227 AA.
AC	
XX	AAM34074;
XX	
DT	27-APR-1998 (first entry)
DE	Hepatitis A virus HM-175 protein sequence.
XX	
KM	HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
KW	Infection; vaccine.
XX	
OS	Hepatitis A virus HM-175.
XX	
FT	Key
FT	Location/Qualifiers
FT	Protein
FT	1..23
FT	/label= VP4
FT	24..245
FT	/label= VP2
FT	246..491
FT	/label= VP3
FT	492..791
FT	/label= VP1
FT	792..980
FT	/label= 2A
FT	981..1087
FT	/label= 2B
FT	1088..1422
FT	/label= 2C
FT	1423..1496
FT	/label= 3A
FT	1497..1519
FT	/label= 3B
FT	1520..1738
FT	/label= 3C
FT	1739..2227
FT	/label= 3D
PN	
XX	M09740166-A2.
PD	
XX	30-OCT-1997.
PF	
XX	18-APR-1997; 97WO-US06506.
PR	
XX	19-APR-1996; 96US-0015642.
PA	(USSH) US SEC DEPT HEALTH.
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PI	
DR	Emerson SU, Purcell RH, Raychaudhuri G;
DR	WPI; 1997-535650/49.
DR	N-PSDB; AAT93023.
PT	
PT	Human attenuated HAV genome containing simian HAV 2C gene - useful as vaccines against HAV infection
PS	
PS	Disclosure; Fig 13A-D; 66pp; English.
CC	This protein sequence is encoded by the human hepatitis A virus
CC	(HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC	HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC	kidney cells. A claimed DNA construct (1) comprises a genome of
CC	HAV, where the genome is a human attenuated HAV genome in which a
CC	region of the 2C gene has been replaced by a corresponding region
CC	from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC	region of the 2C gene from AGM-27 contained in the construct
CC	preferably encodes amino acids 120-328 of the 2C protein, amino
CC	acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC	transcript of (1); (2) a cell transfected with (1) or the RNA

CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
 CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
 CC its RNA transcript, can be used as a vaccine for preventing HAV in
 CC a mammal. (1) or the RNA transcript can also be used to stimulate
 CC the production of protective antibodies in the mammal.

XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 18; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.2e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVQEIKEQ 20
 DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 8

AAB18607 standard; Protein; 2227 AA.

AC AAB18607;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

OS US6113912-A.

PN US6113912-A.

PD 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX MPI; 2000-586464/55.

DR N-PSDB; AAA75476.

DR N-PSDB; AAA75476.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

XX line useful as vaccine for protecting humans against hepatitis A virus

PS infection, has modified genome compared to wild type

XX Disclosure; Fig 6A-K; 72pp; English.

CC The present sequence is derived from a wild type hepatitis A virus

CC (HAV) strain HM-174. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

OY The HAV is able to propagate in MRC-5 cells and retain appropriate

DB attenuation. It is useful as a live vaccine for prophylaxis of

XX hepatitis A in humans and other primates.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVQEIKEQ 20

DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 9

AAB18608
 ID AAB18608 standard; Protein; 2227 AA.

AC AAB18608;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX P-35 virus.

OS Hepatitis A virus.

PN US6113912-A.

PD 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX MPI; 2000-586464/55.

DR N-PSDB; AAA75477.

DR N-PSDB; AAA75477.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

XX line useful as vaccine for protecting humans against hepatitis A virus

PS infection, has modified genome compared to wild type

XX Disclosure; Columns 67-78; 72pp; English.

CC The present sequence is derived from passage 35 of a wild type

CC hepatitis A virus (HAV) strain HM-174. The resulting virus is

CC designated P-35 virus. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

OY The HAV is able to propagate in MRC-5 cells and retain appropriate

DB attenuation. It is useful as a live vaccine for prophylaxis of

XX hepatitis A in humans and other primates.

SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVQEIKEQ 20

DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 10

AAB18609 standard; Protein; 2227 AA.

Query Match 100.0%; Score 96; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVQEIKEQ 20

DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 9

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMK) SMITHKLINE BECHAM BIOLOGICALS.
PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX
DR WPI: 2002-680946/73.
DR N-PSDB: ABS52788.
XX
XX
PS Example 3; Column 67-78; 71pp; English.
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX
CC The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC mutant strain HM-175/77 (pHAV/77) polypeptide.
XX
XX Sequence 2227 AA:
Query Match 100.0%; Score 96; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980
RESULT 13
ABG31729
ID ABG31729 standard; Protein; 2227 AA.
XX
XX ABG31729;
AC
XX
XX 29-NOV-2002 (first entry)
DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX HAV 4380.
XX
XX Hepatitis A virus strain HM-175.
XX
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (SMK) SMITHKLINE BECHAM BIOLOGICALS.
XX
XX Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX WPI: 2002-680946/73.
XX N-PSDB: ABS52789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX
XX Disclosure; Column 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents an attenuated
XX hepatitis A virus 4830 polypeptide.

XX
SQ Sequence 2227 AA:
Query Match 100.0%; Score 96; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980
RESULT 14
AAE19899
ID AAE19899 standard; Protein; 2227 AA.
XX
XX AAE19899;
AC
XX
XX 18-JUN-2002 (first entry)
DE
XX
XX Hepatitis A virus (HAV) protein.
DE
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytoskeletal; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX
XX WO200213855-A2.
XX
XX 21-FEB-2002.
PD
XX
XX 15-AUG-2001; 2001WO-IB01808.
XX
XX 17-AUG-2000; 2000US-225767P.
XX 29-AUG-2000; 2000US-229175P.
XX 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX
XX Salberg M, Hultgren C;
PI
XX
XX WPI: 2002-241837/29.
DR N-PSDB: AAD31766.
XX
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus -
XX
XX
XX Claim 11; Page 82-87; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to
XX a hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is hepatitis A virus (HAV) protein.
XX
XX
XX Sequence 2227 AA:
Query Match 100.0%; Score 96; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980
RESULT 15
ABU08639
ID ABU08639 standard; Protein; 2227 AA.

```

xx AB008639;
xx
xx
xx 03-JUN-2003 (first entry)
xx
xx Wild type human hepatitis A virus strain HM-175.
xx
xx Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
xx vaccine; MRC-5 cell; hepatitis infection.
xx
xx Hepatitis A virus strain HM-175.
xx
xx US2002176869-A1.
xx
xx
xx 28-NOV-2002.
xx
xx
xx 29-APR-2002; 2002US-0135988.
xx
xx 07-JUN-1995; 95US-0475886.
xx PR 31-AUG-2000; 2000US-0653499.
xx PR 18-SEP-1992; 92US-0947338.
xx 17-SEP-1993; 93WO-US08610.
xx 17-APR-1995; 95US-0397232.
xx
xx (FUNK/) FUNKHOUSER A W.
xx PA (EMERSON) EMERSON S U.
xx PA (PURCELL) PURCELL R H.
xx PA (DHONDT) D'HONDT E.
xx
xx Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
xx
xx WPI; 2003-352605/33.
xx DR N-PSTDB; ABX93473.
xx
xx New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
xx useful in vaccines for protecting primates against hepatitis infection
xx and disease -
xx PT
xx PS
xx Disclosure; Fig 6; 70pp; English.
xx
xx The invention describes a live hepatitis A virus (HAV) adapted to growth
xx in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
xx a vaccine for protecting primates against hepatitis infection and
xx CC disease. This is the amino acid sequence of wild type human hepatitis A
xx virus strain HM-175.
xx
xx Sequence 2227 AA;
xx
Query Match 100.0%; Score 96; DB 24; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
1 KINLADRMIGLSGVOEIKQ 20
|||||
661 KINLADRMIGLSGVOEIKQ 980

```

Search completed: October 1, 2003, 09:56:45
Job time : 44.5294 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 seconds
(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVOEIKQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : PIR_76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2227	1	GNMYHM genome polypeptide
2	96	100.0	2227	1	GNMYHM genome polypeptide
3	96	100.0	2227	1	GNMYHM genome polypeptide
4	96	100.0	2227	1	GNMYHM genome polypeptide
5	96	100.0	2227	1	GNMYHM genome polypeptide
6	96	100.0	2227	1	GNMYHM genome polypeptide
7	96	100.0	2227	1	GNMYHM genome polypeptide
8	96	100.0	2227	1	GNMYHM genome polypeptide
9	96	100.0	2227	1	GNMYHM genome polypeptide
10	96	100.0	2227	1	GNMYHM genome polypeptide
11	96	100.0	2227	1	GNMYHM genome polypeptide
12	96	100.0	2227	1	GNMYHM genome polypeptide
13	96	100.0	2227	1	GNMYHM genome polypeptide
14	96	100.0	2227	1	GNMYHM genome polypeptide
15	96	100.0	2227	1	GNMYHM genome polypeptide
16	96	100.0	2227	1	GNMYHM genome polypeptide
17	96	100.0	2227	1	GNMYHM genome polypeptide
18	96	100.0	2227	1	GNMYHM genome polypeptide
19	96	100.0	2227	1	GNMYHM genome polypeptide
20	96	100.0	2227	1	GNMYHM genome polypeptide
21	96	100.0	2227	1	GNMYHM genome polypeptide
22	96	100.0	2227	1	GNMYHM genome polypeptide
23	96	100.0	2227	1	GNMYHM genome polypeptide
24	96	100.0	2227	1	GNMYHM genome polypeptide
25	96	100.0	2227	1	GNMYHM genome polypeptide
26	96	100.0	2227	1	GNMYHM genome polypeptide
27	96	100.0	2227	1	GNMYHM genome polypeptide
28	96	100.0	2227	1	GNMYHM genome polypeptide
29	96	100.0	2227	1	GNMYHM genome polypeptide

30	41	42.7	231	2	B98297	hypothetical prote
31	41	42.7	313	2	D95866	probable transcrip
32	41	42.7	366	1	TDKSM4	monocyte surface g
33	41	42.7	518	1	S44183	thiamine-phosphate
34	41	42.7	605	2	D83007	regulatory protein
35	41	42.7	621	2	A71516	hypothetical prote
36	41	42.7	687	2	JC6329	pepa protein - pas
37	41	42.7	708	2	JC6329	yeast secretory pr
38	41	42.7	1036	2	F83263	probable multitrng
39	41	42.7	1238	2	T40993	protein kinase cck
40	41	42.7	1652	2	I50711	complement C3 prec
41	40	41.7	143	1	HRKW48	heat shock protein
42	40	41.7	208	2	A70122	glucose inhibited
43	40	41.7	247	2	F90656	hypothetical prote
44	40	41.7	247	2	F85507	hypothetical prote
45	40	41.7	332	2	JC4535	urate oxidase (EC

ALIGNMENTS

RESULT 1

GNMYHM genome polypeptide - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Accession: A25981
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
R: Cohen, J.L.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M14707; NID:q329582; PIDN:AAA45465.1; PID:q329583
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 961 KINLADRMGLSGVOEIKQ 980

RESULT 2
GNMYHM genome polypeptide - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R: Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127
 A:Accession: A03903
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <NA>
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
 C:Accession: JS0303
 R:Paul, A.Y.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhard
 A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
 A:Reference number: JS0303; MUID:88045071; PMID:2823500
 A:Accession: JS0303
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <PAD>
 A:Cross-references: EMBL:M20273
 C:Superfamily: hepatitis A virus genome polypeptide
 C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
 F:1-23/Product: coat protein 1A #status predicted <VP1>
 F:24-246/Product: coat protein 1B #status predicted <VP2>
 F:247-491/Product: coat protein 1C #status predicted <VP3>
 F:492-836/Product: coat protein 1D #status predicted <VP1>
 F:837-980/Product: coat protein 2A #status predicted <P2A>
 F:981-1108/Product: core protein 2B #status predicted <P2B>
 F:1109-1436/Product: core protein 2C #status predicted <P2C>
 F:1439-1496/Product: protein 3A #status predicted <P3A>
 F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
 F:1520-1736/Product: cysteine proteinase; protein 3C #status predicted <P3C>
 F:1737-2227/Product: RNA-directed RNA polymerase; protein 3D #status predicted <P3D>

Query Match

Best Local Similarity 100.0%; Score 96; DB 1; Length 2227;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRMGLSGVOEIKQ 20
 ||||||||||||||||

961 KINLADRMGLSGVOEIKQ 980

RESULT 3

GNNTMK
 genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
 NA polymerase (EC 2.7.7.48), protein 3D
 C:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
 C:Accession: A94149; A25914; A94508
 R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Felstone, S.M.; Purcell, F
 Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
 A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
 A:Reference number: A94149; MUID:87175701; PMID:3031686
 A:Accession: A94149
 A:Status: nucleic acid sequence not shown
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <COH>
 A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
 A:Note: submitted to GenBank, August 1987
 C:Superfamily: hepatitis A virus genome polypeptide
 C:Keywords: coat protein; core protein; cysteine proteinase; hydrolyase; nucleotidyltrans
 F:1-245/Product: coat protein 1A #status predicted <P1A>
 F:246-491/Product: coat protein 1B #status predicted <P1B>
 F:492-836/Product: coat protein 1C #status predicted <P1C>
 F:837-980/Product: coat protein 2A #status predicted <P2A>
 F:981-1076/Product: core protein 2B #status predicted <P2B>
 F:1077-1432/Product: core protein 2C #status predicted <P2C>
 F:1423-1484/Product: protein 3A #status predicted <P3A>
 F:1485-1507/Product: protein 3B #status predicted <P3B>
 F:1508-1678/Product: cysteine proteinase; protein 3C #status predicted <P3C>
 F:1679-2227/Product: RNA-directed RNA polymerase; protein 3D #status predicted <P3D>

Query Match

Best Local Similarity 100.0%; Score 96; DB 1; Length 2227;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVOEIKQ 20
 ||||||||||||||||

Db 961 KINLADRMGLSGVOEIKQ 980

RESULT 4

GNNTMB
 genome polypeptide - human hepatitis A virus (strain MBJ)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
 VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus
 A:Note: host Homo sapiens (man)
 C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
 C:Accession: JS0303
 R:Paul, A.Y.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhard
 A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
 A:Reference number: JS0303; MUID:88045071; PMID:2823500
 A:Accession: JS0303
 A:Molecule type: genomic RNA
 A:Residues: 1-2227 <PAD>
 A:Cross-references: EMBL:M20273
 C:Superfamily: hepatitis A virus genome polypeptide
 C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
 F:1-23/Product: coat protein 1A #status predicted <VP1>
 F:24-246/Product: coat protein 1B #status predicted <VP2>
 F:247-491/Product: coat protein 1C #status predicted <VP3>
 F:492-836/Product: coat protein 1D #status predicted <VP1>
 F:837-980/Product: coat protein 2A #status predicted <P2A>
 F:981-1108/Product: core protein 2B #status predicted <P2B>
 F:1109-1436/Product: core protein 2C #status predicted <P2C>
 F:1439-1496/Product: protein 3A #status predicted <P3A>
 F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
 F:1520-1736/Product: cysteine proteinase; protein 3C #status predicted <P3C>
 F:1737-2227/Product: RNA-directed RNA polymerase; protein 3D #status predicted <P3D>

Query Match

Best Local Similarity 100.0%; Score 96; DB 1; Length 2227;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVOEIKQ 20
 ||||||||||||||||

Db 961 KINLADRMGLSGVOEIKQ 980

RESULT 5

GNNTSA
 genome polypeptide - simian hepatitis A virus (strain AGM-27)
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
 C:Species: simian hepatitis A virus
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
 C:Accession: A30470; S04885; S03965
 R:Tzarev, S.A.
 submitted to JIPID, April 1991
 A:Reference number: A30470
 A:Accession: A30470
 A:Molecule type: genomic RNA
 A:Residues: 1-2230 <TSA>
 A:Cross-references: GB:000924; NID:g222597; PIDN:BA00766.1; PID:g222598
 R:Tzarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
 J. Gen. Virol. 72, 1677-1683, 1991
 A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
 A:Reference number: J01080; MUID:91311420; PMID:1649901
 A:Contents: annotation
 A:Note: neither amino acid nor nucleotide sequence is given
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tzarev, S.A.; Sverdlov, E.D.; Chizh
 submitted to the EMBL Data Library, May 1989
 A:Reference number: S04885
 A:Accession: S04885
 A:Molecule type: genomic RNA
 A:Residues: 1750-2164 <BAL>
 A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CAA33490.1; PID:g930268
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tzarev, S.A.; Sverdlov, E.D.; Chizh
 FEBS Lett. 247, 425-428, 1989
 A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
 A:Reference number: S03965; MUID:89232168; PMID:2541023
 A:Accession: S03965
 A:Molecule type: genomic RNA
 A:Residues: 1960-2164 <BAL>
 A:Cross-references: EMBL:X15461
 C:Superfamily: hepatitis A virus genome polypeptide
 C:Keywords: coat protein; core protein; polypeptide
 F:1-27/Product: coat protein 1A #status predicted <C1A>

F:28-249/Product: coat protein 1B #status predicted <C1B>
 F:250-499/Product: coat protein 1C #status predicted <C1C>
 F:496-799/Product: coat protein 1D #status predicted <C1D>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1099-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2230;
 Best Local Similarity 100.0%; Pred. No. 2.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
 |||||
 Db 965 KINLADRMGLSGVOEIKQ 984

RESULT 6

heat shock 16K protein 2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
 C:Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 13-Aug-1999
 C:Accession: B25199
 C:Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.
 J. Biol. Chem. 261, 12006-12015, 1986
 A:Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis
 A:Reference number: A92555; MID:86304344; PMID:3017958
 A:Accession: B25199
 A:Molecule type: DNA
 A:Residues: 1-145 <CON>
 A:Cross-references: GB:M14334; NID:g156338; PIDN:AAA8071.1; PID:g156340
 C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 145;
 Best Local Similarity 55.6%; Pred. No. 2.2;
 Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIK 18
 |||||
 Db 65 KINLADRMGLSGVOEIK 82

RESULT 7

heat shock 16K protein 2 - Caenorhabditis elegans

C:Species: Nippostrongylus brasiliensis
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S33416
 C:Ritvedle, S.; Grig, M.E.; Ingram, L.; Selkirk, M.E.
 submitted to the EMBL Data Library, April 1993
 A:Description: The expression of a small heat shock homologue is developmentally regulated
 A:Reference number: S33416
 A:Accession: S33416
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-172 <TWB>
 A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866
 C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 172;
 Best Local Similarity 45.0%; Pred. No. 2.6;
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
 |||||
 Db 80 KVALDRDLTVBGMDEVKTE 99

RESULT 8

heat shock 16K protein 2 - Caenorhabditis elegans

enoyl-CoA hydratase homolog F11C18.10 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 22-Oct-1999
 C:Accession: T06303
 C:Bevan, M.; Terry, N.; Ardiles, W.; Blythe, C.; Dasseville, R.; De Clerck, R.;
 ewes, H.W.; Mayer, K.F.X.; Schueller, C.
 submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15589
 A:Accession: T06303
 A:Molecule type: DNA
 A:Residues: 1-414 <REV>
 A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:F11C18.10
 A:Experimental source: cultivar Columbia; BAC clone F11C18
 C:Genetics:
 A:Gene: ATSP:F11C18.10
 A:Map position: 4
 A:Introns: 42/3; 75/2; 108/1; 134/3; 162/3; 191/1; 219/3; 260/2; 280/3; 313/3; 347/3;

Query Match 45.8%; Score 44; DB 2; Length 414;
 Best Local Similarity 52.6%; Pred. No. 22;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIK 19
 |||||
 Db 260 KIELDKYKGLDVEIE 278

RESULT 9

GRF-binding membrane protein lepa - Mycobacterium leprae

N:Alternate names: protein B1937_f3_81
 C:Species: Mycobacterium leprae
 C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
 C:Accession: S72609
 C:Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A:Reference number: S72580
 A:Accession: S72609
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-646 <SMI>
 A:Cross-references: EMBL:U00016; NID:g466961; PIDN:AA11717.1; PID:g466991
 C:Genetics:
 A:Gene: lepa
 A:Start codon: GNG
 A:Superfamily: GTP-binding membrane protein lepa; translation elongation factor Tu ho
 C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
 F:45-177/Domain: translation elongation factor Tu homology <ETU>
 F:51-58/Region: nucleotide-binding motif A (P-loop)
 F:174-177/Region: GTP-binding NKXD motif

Query Match 45.8%; Score 44; DB 2; Length 646;
 Best Local Similarity 71.4%; Pred. No. 35;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14
 |||||
 Db 57 KSLADRMGLSGV 70

RESULT 10

GRF-binding membrane protein lepa - Mycobacterium leprae

C:Species: Mycobacterium tuberculosis (strain H37RV)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 02-Feb-2001
 C:Accession: G70683
 C:Cole, S.T.; Broch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 Rajandream, M.A.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd,
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MID:98295987; PMID:9634230

A:Accession: G70683
A:Stetus: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-653 <COL>
A:Cross-references: GB:281368; GB:AL123456; NID:g3261656; PIDN:CAB03723.1; PID:g1655655
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: *lepA*
C:Superfamily: GTP-binding membrane protein *lepA*; translation elongation factor Tu homolog
C:Keywords: GTP binding; nucleotide binding; P-loop
F:55-181/Domain: translation elongation factor Tu homology <ETU>
F:59-66/Region: nucleotide-binding motif A (P-loop)
F:178-181/Region: GTP-binding NKXD motif

	Query Match	45.8%	Score 44;	DB 2;	Length 653;
	Best Local Similarity	71.4%;	Pred. NO. 36;		
Matches	10; Conservative	1;	Mismatches	3;	Indels 0;
Gaps	0;				
QY	1 KINLADRMIGSGV	14			
		1:11			
	65 KSTIADRMLQLRGV	78			

```

RESULT 11
S67595
hypothetical protein YDL060w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein D2544
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002
C:Accession: S67595
R:Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67587
A:Accession: S67595
A:Molecule type: DNA
A:Residues: 1-788 <BL0>
A:Cross-references: EMBL:Z74108; NID:g1431062; PTD:g1431063; GSPDB:GN00004; MIPS:YDL060w
A:Experimental source: strain S286C
C:Genetics:
A:Gene: SGD:TSRL; MIPS:YDL060w
A:Cross-references: SGD:S0002218
A:Map position: 4L

```

Query Match	45.8%	Score 44	DB 2	Length 788
Best Local Similarity	61.5%	Pred. No. 44		
Matches	8	Conservative	3	Mismatches 2
				Indels 0
				Gaps 0
8	MGISGVOEIREQ	20		
159	VFGISGVOEIVDE	171		

RESULT 12
 B24289
 heat shock protein 16-1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 28-Sep-1987 #sequenceRevision 30-Sep-1993 #text_change 29-Oct-1999
 C:Accession: B24289; S31037; T25927; T25930
 R:Russnak, R.H.; Candido, E.P.M.
 M.Ol. Cell. Biol. 5, 1268-1278, 1985
 A:Reference number: A24289; MUID:85295957; PMID:403652
 A:Accession: B24289
 A:Molecule type: DNA
 A:Residues: 1-145 <RUS>
 A:Cross-references: EMBL:K03273; NID:q15633; PIDN:AAA2806.1; PID:q156336
 A:Note: the author translated the codon GAT for residue 17 as Tyr
 R:Kay, R.J.; Russnak, R.H.; Jones, D.; Mathias, C.; Candido, E.P.M.
 Nucleic Acids Res. 15, 3723-3741, 1987
 A:title: Expression of intron-containing C. elegans heat shock genes in mouse cells dem-
 y effect of heat shock on the mammalian splicing apparatus.
 A:Reference number: S31036; MUID:87231065; PMID:3586308
 A:Accession: S31037
 A:Status: translation not shown

A: Molecule type: DNA
A: Residues: 41-63 <RAY>
A: Cross-references: EMBL:M31340; MID:g156343; PIDN:AAA28073.1; PID:g5520668
R: Bradshaw, H.
Submitted to the EMBL Data Library, July 1996
A: Description: The sequence of *C. elegans* cosmid T27E4.
A: Reference number: Z20111

A:Status: preliminary; translated from GB/EMBL/DDdbt
A:Molecule type: DNA
A:Residues: 1-145

A:Cross-references: EMBL:U64837; PIDN:AAB04839.1; GSPDB:GN00023; CESP:hsp-16A
A:Experimental source: strain Bristol N2; clone T27E4
A:Accession: T25930
A:Status: preliminary; translated from GB/EMBL/DDdbt
A:Molecule type: DNA
A:Residues: 1-145

A:Cross-references: EMBL:U64837; PIDN:AAB04842.1; GSPDB:GN00023; CESP:hsp-16A
A:Experimental source: strain Bristol N2; clone T27E4
A:Genetics:
A:Gene: CESP:hsp-16A
A:Map position: 5
A:Introns: 42/3
C:Superfamily: alpha-crystallin

Query Match	44.8%	Score 43	DB 2	Length 145	.
Best Local	Similarity 45.0%	Pred. NO. 10			
Matches 9	Conservative 3	Mismatches 8	Indels 0	Gaps 0	

```
QY      1 KINLADRMGLSGVQEIKEQ 20
        |||| | : ||| :
Db      65 KINLDGHTLSIQGEQELKTE 84
```

RESULT 13
S64331
hypothetical protein YGR030C - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein G4068
C:Species: *Saccharomyces cerevisiae*
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
C:Accession: S64331
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64321
A:Molecule type: DNA
A:Residues: 1-158 <RIE>
A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN0000
A:Experimental source: strain 5286C
C:Genetics:
A:Gene: SGD:POP6; MIPS:YGR030C
A:Cross-references: SGD:S0003262
A:Map position: 7R

Query Match	44.8%	Score 43	DB 2	Length 158
Best Local Similarity	64.3%	Pred. No. 11		
Matches	9	Conservative	2	Mismatches 3
				Indels 0
				Gaps 0
QY	1	KINLADRLGSGV	14	
		: : : :		
Db	65	QINMADRLGLQOV	78	

AC0056 RESULT 14
conserved hypothetical protein YP00454 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
R:Accession: AC0056
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.;
deno-Tarragga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB00001; MUID:21470413; PMID:11586360

A; Accession: AC0056

A;Status: prelimina

A; molecule type: DNA

A;ResIdues: 1-180 <KUR>

A;Residues: 1-180 <KUR>
A:Cross-references: GB:

A;CROSS-references: GB:AL590842; PIDN:CAC89310.1; PID:g15978546; GSPDB:GN00175

C: Genetics:

A;Gene: YP00454

C:Superfamily: Escherichia coli conserved yjx protein

Query Match	44.88;	Score 43;	DB 2;	Length 180;
Best local similarity	47.18;	Pred. No. 13;		

Best Local Similarity 47.1%; Pred. NO. 13;
Matches 8: Conservative 3: Mismatches

Matches	8;	Conservative	3;	Mismatches	6;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

```
QY      4 LADRMGLSGVQEIKEQ 20
          | | | | | : : | |
Db     120 LGDEMAVLGSGISNVKQ 136
```

120 LGDEMAVLGSGISNVKQQ 136

0423

0423

oxygen-independent coproporphyrinogen III oxidase - *Aquifex aeolicus*

C;Species: *Aquifex aeolicus*

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999

C;Accession: H70423

C/Accession: H/0423
R/Decker, G.: Warren, P.V.: Gaasterland, T.: Young, W.G.: Lenox, A.L.: Graham, D.E.: Owsen, J.E.: 1981

Ribeckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Over-

V.	253-358	1008
Nature	307	

Nature 392, 353-358, 1998

A;Title: The comple

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: H70423

A;Status: preliminary;
n;recd: 11/04/23

A: Molecule type: DNA
 A: status: preliminary; nucleic acid sequence not shown; translation not shown

```
A:Molecule type: DNA
A:Residues: 1-370 <AOF>
```

A:ResIdues:

A; Cross-Left
A; Eyed-1m

A; Experimental source: strain VF5

C: Superfamily: oxygen-independent coproporphyrinogen oxidase

c, superfamily: oxygen-independent coproporphyrinogen III dehydratase family

Query Match 44.88: Score 43: DB 2: Length 370: .

QY 6 DRMLGSGVQEI 17

— :: — :: — :: — :: — :: — ::

Search completed: October 1, 2003, 10:04:45
Job time : 14.8235 secs

(
(
(
+
=
(

.
H
A
.
C
B
C
C

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds

(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-46

Sequence: 1 KINADRLGSGVQEKREQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2226	1	POLG_HPAV2
2	96	100.0	2226	1	POLG_HPAV4
3	96	100.0	2226	1	POLG_HPAV8
4	96	100.0	2227	1	POLG_HPAVH
5	96	100.0	2227	1	POLG_HPAVH
6	96	100.0	2227	1	POLG_HPAVH
7	96	100.0	2230	1	POLG_HPAVH
8	48	50.0	622	1	LEPA_STRCO
9	47	49.0	145	1	HS12_CAEEL
10	47	49.0	172	1	HS20_NIPBR
11	44	45.8	646	1	LEPA_MYCLE
12	44	45.8	653	1	LEPA_MYCTU
13	43	44.8	145	1	HS11_CAEEL
14	43	44.8	158	1	POP6_YEAST
15	43	44.8	180	1	Y454_YERPE
16	42.5	44.3	602	1	LEPA_BROME
17	42	43.8	143	1	HS16_CAEEL
18	42	43.8	151	1	SODC_HALRO
19	42	43.8	547	1	YDDB_SALTY
20	42	43.8	621	1	Y708_CHLMU
21	42	43.8	1035	1	CC68_YEAST
22	42	43.8	1250	1	SSD1_YEAST
23	41	42.7	141	1	NIR_METUA
24	41	42.7	155	1	IPR1_SHIFL
25	41	42.7	313	1	CBR1_RHIME
26	41	42.7	313	1	CBR2_RHIME
27	41	42.7	366	1	CD14_MOUSE
28	41	42.7	418	1	ASB6_MOUSE
29	41	42.7	518	1	TH14_SCHPO
30	41	42.7	621	1	Y425_CHLTR
31	41	42.7	708	1	SC10_HUMAN
32	41	42.7	708	1	SC10_RAT
33	41	42.7	1338	1	CEK1_SCHPO

34	40	41.7	143	1	HS17_CAEEL	P02513 caenorhabd
35	40	41.7	208	1	GIDB_BORBU	P53363 borrelia bu
36	40	41.7	316	1	OCDB_HUMAN	Q9UGF7 homo sapien
37	40	41.7	502	1	URIC_BACSB	Q45697 bacillus sp
38	40	41.7	615	1	LEPA_CORGL	Q8nn68 corynebacte
39	40	41.7	659	1	MCHE_ECOTL	Q9enn5 escherichia
40	40	41.7	739	1	RELA_STREQ	Q54089 streptococ
41	40	41.7	819	1	LON_CHLPN	Q92914 chlamydia p
42	40	41.7	1042	1	SVI_BORBU	Q51773 borrelia bu
43	40	41.7	1227	1	CIBB_BACTU	Q05805 bacillus th
44	40	41.7	1228	1	CIBA_BACTK	P05517 bacillus th
45	40	41.7	1229	1	CIBB_BACTU	Q45739 bacillus th

ALIGNMENTS

RESULT 1	ID	POLG_HPAV2	STANDARD:	PRT: 2226 AA.
AC	P26580;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].			
DE	P3D (EC 2.7.7.48)].			
OS	Hepatitis A virus (strain 24a).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Hepatovirus.			
OX	NCBI_TaxID=12094;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91162758; PubMed=1705995;			
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,			
RA	Cromeans T., Jansen R.W.,			
RT	"Antigenic and genetic variation in cytopathic hepatitis A virus			
RT	variants arising during persistent infection: evidence for genetic			
RT	recombination."			
RL	J. Virol. 65:2056-2065(1991).			
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +			
CC	(RNA)(N).			
CC	-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,			
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,			
CC	VP3, AND VP4.			
CC	-1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sb.ch/announce/			
CC	or send an email to license@isb-sb.ch).			
CC	-----			
DR	EMBL: M59810; AAA5468.1; -			
DR	MEROPS: C03.005; -			
DR	InterPro: IPR004004; Calic1_pol_hel.			
DR	InterPro: IPR000605; RNA_helicase.			
DR	InterPro: IPR007095; RNA_pol_DS_PS.			
DR	InterPro: IPR001205; RNA_pol_P3D.			
DR	InterPro: IPR007094; RNA_pol_PSVIR.			
DR	Pfam: PF00680; RNA_dep_RNA_pol.1.			
DR	Pfam: PF00910; RNA_helicase.1.			
DR	PRINTS: PR00918; CALICVIRUSNS.			
KW	polyprotein; Coat protein; Core protein; Transferrase;			
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.			
FT	CHAIN 1 23			
FT	CHAIN 1 23			
FT	CHAIN 24 245			
FT	CHAIN 246 491			
FT	CHAIN 492 794			
FT	CHAIN 795 900			
FT	CHAIN CORE PROTEIN P2A.			

FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MM; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 2

POLG_HPVA4 STANDARD; PRT; 2226 AA.

P26581;

01-AUG-1992 (Rel. 23, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins

P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

P2D (PC 2.7.7.4b)].

05 Hepatitis A virus (strain 43c).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_Taxid=12095;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-9162736; PubMed-1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,

RT "Antigenic and genetic variation in cytopathic hepatitis A virus

RT variants arising during persistent infection: evidence for genetic

RT recombination".

RL J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC [RNA](N).

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,

CC VP3, AND VP4.

CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

CC EMBL, M59809; AAA45469.1; -.

DR MEROPS; C03.005; -.

DR InterPro: IPR000605; RNA_helicase.

DR InterPro: IPR007095; RNA_pol_DS_PS.

DR InterPro: IPR001205; RNA_pol_P3D.

DR InterPro: IPR007094; RNA_pol_PSVir.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).

FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).

FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).

FT CHAIN 795 900 CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 1088 1422 CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MM; 403B4CA80B95F75 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 3

POLG_HPVA8 STANDARD; PRT; 2226 AA.

P26582;

01-AUG-1992 (Rel. 23, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)

Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins

P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase

P2D (PC 2.7.7.4b)].

05 Hepatitis A virus (strain 18f).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_Taxid=12096;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE-9162736; PubMed-1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,

RT "Antigenic and genetic variation in cytopathic hepatitis A virus

RT variants arising during persistent infection: evidence for genetic

RT recombination".

RL J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC [RNA](N).

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,

CC VP3, AND VP4.

CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).

CC EMBL, M59808; AAA45467.1; -.

DR PDB; 1QAV; 15-MAY-00.

DR MEROPS; C03.005; -.

DR InterPro: IPR000605; RNA_helicase.

DR InterPro: IPR007095; RNA_pol_DS_PS.

DR InterPro: IPR001205; RNA_pol_P3D.

DR InterPro: IPR007094; RNA_pol_PSVir.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.

DR Pfam; PF00910; RNA_helicase; 1.

KW Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.

FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).

FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).

FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).

FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).

FT CHAIN 795 900 CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 1088 1422 CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.

RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02990; AAA45472.1; -.
DR PIR: A03903; GNNTHR.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol_1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CR664;
Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KINLADRLGLSGVQEIKEQ 20
|||||
961 KINLADRLGLSGVQEIKEQ 980

RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MB8).";
RL Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20273; AAA45474.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol_1.
DR Pfam: PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983E2A7C86449 CR664;
Query Match 100.0%; Score 96; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KINLADRLGLSGVQEIKEQ 20
|||||
961 KINLADRLGLSGVQEIKEQ 980

RESULT 7
POLG_HPAVMS STANDARD: PRT: 2230 AA.
ID POLG_HPAVMS
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome

us-09-171-432a-46.rsp

CC Streptomyces: Streptomycetaceae; Streptomycetes.
OX NCBI_taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000931;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.N., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Hubbard C.H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Ruter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1 SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isdb-sib.ch/announce/
CC or send an email to license@isdb.sib.ch).

CC -----

DR EMBL; AL939113; CAB66240.1; -
DR HSSP; P13551; 2EFG.
DR HAMAP; MF_00071; ?; 1.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR000640; EFG_C.
DR InterPro: IPR004161; EFFU_D2.
DR InterPro: IPR006297; LepA.
DR InterPro: IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNFCT.
DR TRGFAMS; TIGR01393; LepA; 1.
DR TRGFAMS; TIGR00231; small_gtp; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW GTP-binding; Complete proteome.

FT NP_BIND 26 GTP (BY SIMILARITY).
FT FT 94 GTP (BY SIMILARITY).
FT NP_BIND 148 151 GTP (BY SIMILARITY).
SQ SEQUENCE 622 AA; 68378 MW; 83FF5C76FA2AB0C7C CNC64;

Query Match 50.0%; Score 48; DB 1; Length 622;
Best Local Similarity 57.9%; Pred. NO. 4.1;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KINLADRMGLGSGVOEIKE 19
DB | ||||| :| :: 50
32 KSTLDADMQLTLGVVEGRQ 50

RESULT 9
AC HS12_CAEBL STANDARD; PRN: 145 AA.
ID P06582;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Heat shock protein HSP16-2.
GN HSP16-2 OR Y46H3A.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoea; Chromadorea; Rhabditidae; Rhabditoides;
OC Rhabditidae; Pelodierinae; Caenorhabdilis.
OX NCBI_taxid=6239;
NN [1]

CC Streptomyces: Streptomycetaceae; Streptomycetes.
OX NCBI_taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thompson N.R., James K.D., Harris D.E., Quall M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.R., Collins M.,
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
RA Hubbard C.H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Ruter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).

DR EMBL; AL939113; CAB66240.1; -
DR HSSP; P13551; 2EFG.
DR HAMAP; MF_00071; -; 1.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR006297; LepA.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR01393; LepA; 1.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW GTP-binding; Complete proteome.
FT NP_BIND 26 GTP (BY SIMILARITY).
FT NP_BIND 94 GTP (BY SIMILARITY).
FT NP_BIND 148 151 GTP (BY SIMILARITY).
SQ SEQUENCE 622 AA; 68378 MW; 83F5C76F82A80C7C CAC64;

Query Match 50.0%; Score 48; DB 1; Length 622;
Best Local Similarity 57.9%; Pred. NO. 4.1;
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 KINLADRMILGSGVOEIKE 19
DB 32 KSTLADRMILGTVGEORQ 50
| | | | | | | | | | | | : : :

RESULT 9
HS12_CAEEL
ID HS12_CAEEL STANDARD; PRT; 145 AA.
AC P06582;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Heat shock protein HSP16-2.
DE HSP16-2 OR Y46H3A.3.
GN Caenorhabditis elegans.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
OC Rhabditidae; Peloiderinae; Caenorhabditis.
OX NCBI_taxid=6239;
NN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE-6304344; PubMed-3017958;
RA Jones D., Russnak R.H., Kay R.J., Candido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
  Caenorhabditis elegans that is flanked by repetitive elements.";
RL J. Biol. Chem. 261:12006-12015(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Stoneking T., Wohlmann P., Lennox S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
  FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14334; AAA28071.1; -
DR EMBL; AC006774; AAF60615.1; -
DR PIR; B25199; B25199.
DR WormPep; Y46H3A.3; CE22002.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
DR Heat shock; Multigene family.
KW SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;
SQ
Query Match 49.0%; Score 47; DB 1; Length 145;
Best Local Similarity 55.6%; Pred. No. 1.3;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
Oy 1 KINLADRLGLSGVQEI 18
Db 65 KINLDRGLSLGQVQELK 82
RESULT 10
HSP20_NIPBR STANDARD; PRT; 172 AA.
ID HS20_NIPBR 007160;
01-OCT-1994 (Rel. 30, Created)
01-OCT-1994 (Rel. 30, Last sequence update)
01-OCT-1994 (Rel. 30, Last annotation update)
DE Heat shock protein homolog (HSP20).
HSP20.
GN Nipostrogyllus brasiliensis.
OS Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloidea;
OC Trichostrongyloidea; Heligmonellidae; Nipostrogyllinae;
OC Nipostrogyllus.
OC NCBI_TaxID=27835;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94081860; PubMed-8259127;
RA Tweedle S., Grigov M.E., Ingram L., Selkirk M.E.;
RT "The expression of a small heat shock protein homologue is
  developmentally regulated in Nipostrogyllus brasiliensis.";
RL Mol. Biochem. Parasitol. 61:149-154(1993).
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
  FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71663; CA50655.1; -
DR PIR; S33416; S33416.
DR InterPro; IPR001436; Crystallin_alpha.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF00011; HSP20; 1.
DR PRINTS; PR00299; ACRYSTALLIN.
DR PROSITE; PS01031; HSP20; 1.
DR Heat shock.
KW SEQUENCE 172 AA; 20227 MW; 2CDA711CE60B1C0 CRC64;
SQ
Query Match 49.0%; Score 47; DB 1; Length 172;
Best Local Similarity 45.0%; Pred. No. 1.6;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
Oy 1 KINLADRLGLSGVQEI 20
Db 80 KVLDDRLDRLVEGMEVQTE 99
RESULT 11
LEPA_MYCLE STANDARD; PRT; 646 AA.
ID LEPA_MYCLE P53530;
AC P53530;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR ML0611 OR B1937_F3_81.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OC NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE-21128732; PubMed-11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Felkell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagers K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- MISCELLANEOUS: THE OVERPRODUCTION OF THIS PROTEIN IS LETHAL TO
  M. LEPAE.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
  LEPA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00016; AAA17177.1; -
DR EMBL; AL583919; CAC30119.1; -
DR PIR; S72609; S72609.
DR HSP; P13551; 2EPG.
DR Leproma; ML0611; -

```

DR HAMAP: MF_00071; -- 1.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR000640; EFG_C.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR006297; LepA.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00679; EFG_C.1.
 DR Pfam: PF00009; GTP-EFTU.1.
 DR Pfam: PF03144; GTP-EFTU_D2.1.
 DR PRINTS: PR00315; ELONGATNFCF.
 DR TIGRPFAMS: TIGR01393; LepA.1.
 DR TIGRPFAMS: TIGR00231; small_GTP.1.
 DR PROSITE: PS00301; EFACITOR_GTP.1.
 DR GTP-binding; Complete proteome.
 KW NP_BIND 51 58 GTP (BY SIMILARITY).
 FT NP_BIND 120 124 GTP (BY SIMILARITY).
 FT NP_BIND 174 177 GTP (BY SIMILARITY).
 SO SEQUENCE 646 AA; 71329 MW; 256DC1AE894C4A3 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 646;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14
 | | | | | | | | | | | | | | | | | |
 Db 57 KSTLADRMQLTGV 70

RESULT 12

LEPA_MYCTU STANDARD; PRT; 653 AA.
 AC P71739;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE GTP-binding protein lepa.
 GN LEPA OR RV2404C OR WT2476 OR MTCY253.16.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinomycetales;
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid-1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekla F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 Hornsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulton J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey O.,
 Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gali J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: Z81368; CAB03723.1; --
 DR EMBL: AE007086; AAK46772.1; --
 DR PIR: G70683; G70683.
 DR HSSP: P1351; 2EFG.
 DR TIGR: WT2476; --
 DR Tuberculist: RV2404C; --
 DR HAMAP: MF_00071; -- 1.
 DR InterPro: IPR000795; EF_GTPbind.
 DR InterPro: IPR000640; EFG_C.
 DR InterPro: IPR004161; EFTU_D2.
 DR InterPro: IPR006297; LepA.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00679; EFG_C.1.
 DR Pfam: PF00009; GTP-EFTU.1.
 DR Pfam: PF03144; GTP-EFTU_D2.1.
 DR PRINTS: PR00315; ELONGATNFCF.
 DR TIGRPFAMS: TIGR01393; LepA.1.
 DR TIGRPFAMS: TIGR00231; small_GTP.1.
 DR PROSITE: PS00301; EFACITOR_GTP.1.
 DR GTP-binding; Complete proteome.
 KW NP_BIND 59 66 GTP (BY SIMILARITY).
 FT NP_BIND 124 128 GTP (BY SIMILARITY).
 FT NP_BIND 178 181 GTP (BY SIMILARITY).
 SO SEQUENCE 653 AA; 72395 MW; DA4AFB10E6C25755 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 653;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14
 | | | | | | | | | | | | | | | | | |
 Db 65 KSTLADRMQLTGV 78

RESULT 13

HS11_CAEEL STANDARD; PRT; 145 AA.
 ID HS11_CAEEL
 AC P34696;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Heat shock protein HSP16-1.
 GN (HSP16-1A OR T27E4.2) AND (HSP16-1B OR T27E4.8).
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_Taxid-6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85295957; PubMed-4033652;
 RA Russnak R.H., Candido E.P.M.;
 RT "Locus encoding a family of small heat shock genes in Caenorhabditis
 RT elegans: two genes duplicated to form a 3.8-kilobase inverted
 RT repeat.";
 RL Mol. Cell. Biol. 5:1268-1278(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Bristol N2;
 RA Bradehaw H.;
 RT Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 41-63 FROM N.A.
 RX MEDLINE-87231065; PubMed-3588308;
 RA Kay R.J., Russnak R.H., Jones D., Mathias C., Candido E.P.M.;
 RT "Expression of iron-containing C. elegans heat shock genes in mouse
 RT cells demonstrates divergence of 3' splice site recognition sequences
 RT between nematodes and vertebrates, and an inhibitory effect of heat
 RT shock on the mammalian splicing apparatus.";


```

RL Nucleic Acids Res. 15:3723-3741(1987).
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
CC DR EMBL; K03273; AAA28068.1; -
CC DR EMBL; U64837; AAB04842.1; -
CC DR EMBL; U64837; AAB04839.1; -
CC DR PIR; B24289; B24289.
CC DR WormPep; T27E4.2; C614249.
CC DR WormPep; T27E4.8; C614249.
CC DR InterPro; IPR001436; CrystalIn_alpha.
CC DR InterPro; IPR002068; Hsp20.
CC Pfam; PF00011; HSP20.1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
CC KW Heat shock; Multigene family.
CC SO SEQUENCE 145 AA; 16253 MW; 06C36A1F06D15A11 CRC64;
CC -----
OY Query Match 44.8%; Score 43; DB 1; Length 145;
Db Best Local Similarity 45.0%; Pred. No. 5.9;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
1 KINLADRLGLSGVOEIKQ 20
1111 1 1 1 1 1 1
65 KINLDGHTLSIQEDELKTE 84
-----
RESULT 14
POP6_YEAST
ID POP6_YEAST STANDARD: PRT: 158 AA.
AC P53218.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonucleases P/MRP protein subunit POP6 (EC 3.1.26.5) (RNases P/MRP
DE 18.2 kDa subunit) (RNA processing protein POP6).
GN POP6 OR YGR030C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NC C1 TaxID=4932;
[1]
SEQUENCE FROM N.A.
RC STRAIN=5288C;
RX MEDLINE=97435481; PubMed=9290212;
RA Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RP "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RP chromosome VII."
RL Yeast 13:1077-1090(1997).
CC -1- FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT
CC GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.
CC ALSO A COMPONENT OF RNASE MRP.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -1- SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P
CC RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOIETY AND AT LEAST
CC 8 PROTEIN SUBUNITS; POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND
CC RPP1.
CC -1- SUBCELLULAR LOCATION: Nuclear (potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
CC DR EMBL; Z72815; CAA97018.1; -.
CC DR PIR; S64321; S64321.
CC DR SGD; S0003262; POP6.
CC DR GO; GO:00001172; C:ribonuclease mitochondrial RNA processing c. . .; IDA.
CC DR GO; GO:0005655; C:ribonuclease P complex; IDA.
CC DR GO; GO:0000171; F:ribonuclease MRP activity; IDA.
CC DR GO; GO:0004526; F:ribonuclease P activity; IDA.
CC DR Hydroxylase; Nuclear protein; tRNA processing; Coiled coll.
CC FT DOMAIN 51 71 COILED COIL (POTENTIAL).
CC SO SEQUENCE 158 AA; 18210 MW; 6c27a73f3ad521181 CnC64;
CC
CC Query Match 44.8%; Score 43; DB 1; Length 158;
CC Best Local Similarity 64.3%; Pred. No. 6.5;
CC Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC
CC Oy 1 KINADRLGSGV 14
CC Db 65 QINMADRLSLGQV 78
CC
CC RESULT 15
CC Y454_YERPE STANDARD; PRT; 180 AA.
CC AC Q82IP1;
CC DT 15-SEP-2003 (Rel. 42, Created)
CC DT 15-SEP-2003 (Rel. 42, Last sequence update)
CC DE 15-SEP-2003 (Rel. 42, Last annotation update)
CC DT Hypothetical UPF0244 protein YP00454/Y3725.
CC GN YP00454 OR Y3725.
CC OS Yersinia pestis.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
CC OC Enterobacteriaceae; Yersinia.
CC OX NCBI_TaxID=632;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=CO-92 / Blovor Orientalis;
CC RC MEDLINE=21470413; PubMed=11586360;
CC RX Pakkhil J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
CC RA Prentice M.B., Sebahlia M., James K.D., Churcher C., Mungall K.L.,
CC RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
CC RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
CC RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
CC RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
CC RA Slimmons M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
CC RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
CC RL Nature 413:523-527(2001).
CC [2]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=KIMS / Blovor Mediaevalis;
CC RC MEDLINE=22137863; PubMed=12142430;
CC RX Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,
CC RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
CC RA Fehreston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
CC RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
CC RA Perry R.D.;
CC RT "Genome sequence of Yersinia pestis KIM.";
CC RL J. Bacteriol. 184:4601-4611(2002).
CC CC -1- SIMILARITY: Belongs to the UPF0244 family.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC -----
CC CC EMBL; AJ141412; CAC89310.1; -.
CC CC EMBL; AE013976; ABAM6723.1; -.
CC CC PIR; AC0056; AC0056.
CC

```


DR HAMAP: MF_00648: -: 1.
DR InterPro: IPR002786; DUF84.
DR Pfam: PF01931; DUF84; 1.
DR ProDom: PD016591; DUF84; 1.
DR TIGRFAMs: TIGR00258; TIGR00258; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 180 AA; 19694 MW; F95E9653D710828D CRC64;

Query Match 44.8%; Score 43; DB 1; Length 180;
Best Local Similarity 47.1%; Pred. No. 7.4;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 LADRMGLSGVQDEIKQ 20
| | | | | : | : |
DB 120 LGDEMAVLGSGISNVKQ 136

Search completed: October 1, 2003, 09:57:54
Job time : 8.58824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432a-46
Perfect score: 96
Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriophage:*
17: sp_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	251	12 Q9ENP8	Q9enp8 hepatitis a
2	96	100.0	251	12 Q9ENN9	Q9enn9 hepatitis a
3	96	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
4	96	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
5	96	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
6	96	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
7	96	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
8	96	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
9	96	100.0	251	12 Q9ENP4	Q9enp4 hepatitis a
10	96	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
11	96	100.0	251	12 Q9ENN4	Q9enn4 hepatitis a
12	96	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
13	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
14	96	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
15	96	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
16	96	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a

17	96	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
18	96	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
19	96	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
20	96	100.0	251	12 Q9ENQ8	Q9enq8 hepatitis a
21	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
22	96	100.0	251	12 Q9ENQ0	Q9enq0 hepatitis a
23	96	100.0	1124	12 Q84780	Q84780 hepatitis a
24	96	100.0	1161	12 Q05794	Q05794 hepatitis a
25	96	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
26	96	100.0	2218	12 Q67824	Q67824 hepatitis a
27	96	100.0	2218	12 Q67817	Q67817 hepatitis a
28	96	100.0	2225	12 Q9DLJ2	Q9dlj2 hepatitis a
29	96	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
30	96	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
31	96	100.0	2227	12 Q67825	Q67825 hepatitis a
32	96	100.0	2227	12 Q9WMA9	Q9wma9 hepatitis a
33	96	100.0	2227	12 Q80V03	Q80v03 hepatitis a
34	96	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
35	96	100.0	2227	12 Q67826	Q67826 hepatitis a
36	96	100.0	2227	12 Q8V0N6	Q8v0n6 hepatitis a
37	96	100.0	2227	12 Q91FH5	Q91fh5 hepatitis a
38	96	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
39	96	95.8	251	12 Q9ENP0	Q9enp0 hepatitis a
40	96	95.8	251	12 Q9ENN5	Q9enn5 hepatitis a
41	96	95.8	251	12 Q9ENN7	Q9enn7 hepatitis a
42	96	94.8	2225	12 Q9DWR1	Q9dwr1 hepatitis a
43	91	94.8	2227	12 Q8OR16	Q8or16 hepatitis a
44	88	91.7	251	12 Q9ENR0	Q9enr0 hepatitis a
45	88	91.7	251	12 Q9ENN8	Q9enn8 hepatitis a

ALIGNMENTS

RESULT 1

ID Q9ENP8 PRELIMINARY; PRT; 251 AA.
AC Q9ENP8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A304;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DDJ databases.
DR EMBL: AB047665; BAB12173.1; -.
FT NON_TER
FT NON_TER
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20

DB 125 KINLADRMGLSGVQEIKEQ 144

RESULT 2

ID Q9ENN9 PRELIMINARY; PRT; 251 AA.
AC Q9ENN9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatitis A virus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A68;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047674; BAB12182.1; -
FT NON_TER 1
FT SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRMGLSGVOEIKQ 20
125 KINLADRMGLSGVOEIKQ 144

RESULT 3
Q9ENQ4 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ4
AC Q9ENQ4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatitis A virus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A201;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047659; BAB12167.1; -
FT NON_TER 1
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRMGLSGVOEIKQ 20
125 KINLADRMGLSGVOEIKQ 144

RESULT 4
Q9ENQ2 PRELIMINARY; PRT; 251 AA.
ID Q9ENQ2
AC Q9ENQ2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatitis A virus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A205;
RA Fujiwara K.;

RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047661; BAB12169.1; -
FT NON_TER 1
FT SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRMGLSGVOEIKQ 20
125 KINLADRMGLSGVOEIKQ 144

RESULT 5
Q9ENR2 PRELIMINARY; PRT; 251 AA.
ID Q9ENR2
AC Q9ENR2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatitis A virus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A9;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047681; BAB12189.1; -
FT NON_TER 1
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRMGLSGVOEIKQ 20
125 KINLADRMGLSGVOEIKQ 144

RESULT 6
Q9ENP2 PRELIMINARY; PRT; 251 AA.
ID Q9ENP2
AC Q9ENP2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatitis A virus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A5;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB047671; BAB12179.1; -
FT NON_TER 1
FT SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
|||||
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 7
OGEN06 PRELIMINARY; PRT; 251 AA.

AC OGEN06: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;

OX NCB1_TaxID=12092;
(1)
SEQUENCE FROM N.A.

RA Fujikura K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.

FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
|||||
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 8
OGENR1 PRELIMINARY; PRT; 251 AA.

AC OGENR1: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;

OX NCB1_TaxID=12092;
(1)
SEQUENCE FROM N.A.

RA Fujikura K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -.

FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
|||||
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 9
OGENP1

ID OGENP1 PRELIMINARY; PRT; 251 AA.

AC OGENP1: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;

OX NCB1_TaxID=12092;
(1)
SEQUENCE FROM N.A.

RA Fujikura K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -.

FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
|||||
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 10
OGENP4 PRELIMINARY; PRT; 251 AA.

AC OGENP4: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;

OX NCB1_TaxID=12092;
(1)
SEQUENCE FROM N.A.

RA Fujikura K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047669; BAB12177.1; -.

FT NON_TER 1 1
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20
|||||
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 11
OGEN09 PRELIMINARY; PRT; 251 AA.

AC OGEN09: 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

```

ON NCBI_TaxID=12092;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A159;
RA Fujiwara K.;
RL "hepatitis A virus.";
RR Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
FT EMBL; AB047654; BABI2162.1; -.
FF NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482862F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 KINLADRMIGLSGVOEIKQ 20
Db 125 KINLADRMIGLSGVOEIKQ 144
|||||
|||

RESULT 12
OGENN4 PRELIMINARY; PRT; 251 AA.
AC OGENN4;
AC OGENN4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
ON [1]
RN SEQUENCE FROM N.A.
RP STRAIN=A77;
RC Fujiwara K.;
RA "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RR EMBL; AB047679; BABI2187.1; -.
FF NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KINLADRMIGLSGVOEIKQ 20
|||||
|||

125 KINLADRMIGLSGVOEIKQ 144
|||||
|||

RESULT 13
OGENP5 PRELIMINARY; PRT; 251 AA.
AC OGENP5;
AC OGENP5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
ON [1]
RN SEQUENCE FROM N.A.
RP STRAIN=A407;
RC Fujiwara K.;
RA "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RR EMBL; AB047668; BABI2176.1; -.
FF NON_TER 1

```

FT	NON_TER	251	251	8334EF179C757A6D	CRC64:
5Q	SEQUENCE	251 AA:	28614 MW;		
	Query Match				
	Best Local Similarity	100.0%;	Score 96;	DB 12;	Length 251;
	Matches	20; Conservative	0;	Mismatches	0; Indels
0Y					
	1 KINLADRLGLSGVOEIKQ	20			
DB					
	125 KINLADRLGLSGVOEIKQ	144			
	RESULT 14				
09ENQ3					
ID	09ENQ3	PRELIMINARY;	PRT;	251 AA.	
AC	09ENQ3;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment).				
OS	Hepatitis A virus.				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OX	Hepatitis A virus.				
NCBI_TaxID=12092;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-A204;				
RA	Fujiiwara K.;				
RT	"hepatitis A virus.";				
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB047660; BAB12168.1; -.				
FT	NON_TER	1			
FT	NON_TER	251			
5Q	SEQUENCE	251 AA;	28693 MW;	C32AD5651506751A	CRC64;
	Query Match				
	Best Local Similarity	100.0%;	Score 96;	DB 12;	Length 251;
	Matches	20; Conservative	0;	Mismatches	0; Indels
0Y					
	1 KINLADRLGLSGVOEIKQ	20			
DB					
	125 KINLADRLGLSGVOEIKQ	144			
	RESULT 15				
09ENP7					
ID	09ENP7	PRELIMINARY;	PRT;	251 AA.	
AC	09ENP7;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment).				
OS	Hepatitis A virus.				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OX	Hepatitis A virus.				
NCBI_TaxID=12092;					
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-A306;				
RA	Fujiiwara K.;				
RT	"hepatitis A virus.";				
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; AB047666; BAB12174.1; -.				
FT	NON_TER	1			
FT	NON_TER	251			
5Q	SEQUENCE	251 AA;	28720 MW;	C3342482882F19CA	CRC64;
	Query Match				
	Best Local Similarity	100.0%;	Score 96;	DB 12;	Length 251;
	Matches	20; Conservative	0;	Mismatches	0; Indels
0Y					
	1 KINLADRLGLSGVOEIKQ	20			

Db 125 KINLADRMIGSGVOEIKEQ 144

Search completed: October 1, 2003, 10:02:41
Job time : 34.1765 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds
(without alignments)
66,600 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVOEIKQ 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2227	3	US-08-475-886-2
2	96	100.0	2227	3	US-08-475-886-4
3	96	100.0	2227	3	US-08-475-886-6
4	96	100.0	2227	3	US-08-397-232-2
5	96	100.0	2227	3	US-08-397-232-4
6	96	100.0	2227	3	US-09-171-387-2
7	96	100.0	2227	4	US-09-653-499-2
8	96	100.0	2227	4	US-09-653-499-4
9	96	100.0	2227	4	US-09-653-499-6
10	96	100.0	2227	4	US-09-653-499-6
11	42	43.8	1250	1	US-08-441-139-9
12	41	42.7	366	3	US-08-746-883-6
13	41	42.7	677	4	US-09-252-991A-32924
14	41	42.7	708	4	US-09-643-597-369
15	40.5	42.2	295	4	US-09-328-352-5192
16	40	41.7	331	4	US-09-634-238-247
17	40	41.7	332	1	US-08-469-649-2
18	40	41.7	332	4	US-09-347-878-60
19	40	41.7	440	4	US-09-198-452A-44
20	40	41.7	638	2	US-08-846-762-95
21	40	41.7	643	3	US-09-178-252-25
22	40	41.7	1186	3	US-09-178-252-23
23	40	41.7	1207	1	US-07-951-715A-7
24	40	41.7	1207	2	US-08-459-448A-7
25	40	41.7	1207	3	US-08-459-595A-7
26	40	41.7	1207	3	US-08-459-504B-7
27	40	41.7	1207	3	US-08-459-444-7
					Sequence 2, App11
					Sequence 4, App11
					Sequence 6, App11
					Sequence 2, App11
					Sequence 4, App11
					Sequence 2, App11
					Sequence 4, App11
					Sequence 2, App11
					Sequence 4, App11
					Sequence 6, App11
					Sequence 9, App11
					Sequence 6, App11
					Sequence 32924, A
					Sequence 369, App
					Sequence 5192, App
					Sequence 247, App11
					Sequence 60, App11
					Sequence 44, App1
					Sequence 95, App1
					Sequence 25, App1
					Sequence 23, App1
					Sequence 7, App11
					Sequence 7, App11
					Sequence 7, App11
					Sequence 7, App11
					Sequence 7, App11
					Sequence 8, App11

28	40	41.7	1207	4	US-09-547-422-7	Sequence 7, App11
29	40	41.7	1227	1	US-08-448-170-8	Sequence 8, App11
30	40	41.7	1227	3	US-09-053-549-2	Sequence 2, App11
31	40	41.7	1227	3	US-08-961-803-9	Sequence 9, App11
32	40	41.7	1229	1	US-08-100-709-4	Sequence 4, App11
33	40	41.7	1229	1	US-08-176-865-4	Sequence 4, App11
34	40	41.7	1229	1	US-08-474-038-4	Sequence 4, App11
35	40	41.7	1229	2	US-08-779-046-4	Sequence 4, App11
36	40	41.7	1297	2	US-08-881-340-4	Sequence 4, App11
37	40	41.7	1297	4	US-09-107-532A-4552	Sequence 4552, App
38	40	41.7	1388	2	US-08-685-576-1	Sequence 1, App11
39	39	40.6	310	4	US-09-724-623-122	Sequence 122, App
40	39	40.6	318	4	US-09-134-001C-4689	Sequence 4689, App
41	39	40.6	439	4	US-09-252-991A-20350	Sequence 20350, A
42	39	40.6	528	3	US-08-904-871-5	Sequence 5, App11
43	39	40.6	593	4	US-09-252-991A-25274	Sequence 25274, A
44	39	40.6	748	3	US-08-904-871-6	Sequence 6, App11
45	39	40.6	748	3	US-08-904-871-13	Sequence 13, App11

ALIGNMENTS

```
RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 2026426JUS2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match          100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVOEIKQ 20
Db       961 KINLADRMGLSGVOEIKQ 980

RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 2026426JUS2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
```

;; EARLIER FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 3

US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912

;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT FILING DATE: 1995-06-07
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: 08/397,232
;; EARLIER FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 4

US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110

;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US1
;; CURRENT APPLICATION NUMBER: US/08/397,232A
;; CURRENT FILING DATE: 1995-04-17
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: PCT/US93/08610
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227

;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 5

US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110

;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US1
;; CURRENT APPLICATION NUMBER: US/08/397,232A
;; CURRENT FILING DATE: 1995-04-17
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: PCT/US93/08610
;; EARLIER FILING DATE: 1993-09-17
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20
Db 961 KINLADRLGLSGVQEIKEQ 980

RESULT 6

US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734

;; GENERAL INFORMATION:
;; APPLICANT: RAYCHAUDHURI, GOPA;
;; EMERSON, SUZANNE, U.;
;; PURCELL, ROBERT, H.;
;; TITLE OF INVENTION: SIMIAN-HUMAN HAV
;; HAVING A CHIMERIC 2C PROTEIN
;; NUMBER OF SEQUENCES: 3
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: MORGAN & FINNEGAN, L.L.P.
;; STREET: 345 PARK AVENUE
;; CITY: NEW YORK
;; STATE: NEW YORK
;; COUNTRY: USA
;; ZIP: 10154
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: FLOPPY DISK
;; COMPUTER: IBM PC COMPATIBLE
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: MICROSOFT WORD 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/171,387
;; FILING DATE: 24-Mar-1999

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229051
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 8
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U

APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 9
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 96; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 10
US-08-441-139-9
Sequence 9, Application US/08441139
Patent No. 5773245
GENERAL INFORMATION:
APPLICANT: WITTRUP, DR. KARL D.
APPLICANT: RODINSON, ANNE S.
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,139
FILING DATE: 15-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,997
FILING DATE: 06-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, Frank S.
REGISTRATION NUMBER: 31,346
TELEPHONE: 516-742-4343
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-441-139-9

Query Match 43.8%; Score 42; DB 1; Length 1250;
Best Local Similarity 42.1%; Pred. NO. 57;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 INLADRMGLSGVOEIKR 20
DB 871 INLADRMGLSGVOEIKR 889

RESULT 11
US-08-746-883-6
Sequence 6, Application US/08746883
Patent No. 6093693
GENERAL INFORMATION:
APPLICANT: Julius, Michael H., Philipp, Dominik,
APPLICANT: Alizadeh-Khavi, Kamel
TITLE OF INVENTION: B Cell Activation
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Blake, Cassels & Graydon
STREET: Box 25, Commerce Court West
CITY: Toronto
COUNTRY: Canada
ZIP: M5L 1A9
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage
OPERATING SYSTEM: MS-DOS 5.1
SOFTWARE: WORD PERFECT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/746,883
FILING DATE: NO. 6093693ember 18, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunt, John C.
REGISTRATION NUMBER: 36,424
REFERENCE/DOCKET NUMBER: 47841/00008

TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 863-4344
TELEFAX: (416) 863-2653
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 366 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-746-883-6

Query Match 42.7%; Score 41; DB 3; Length 366;
Best Local Similarity 63.6%; Pred. NO. 19;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 7 RMLGSGVOEI 17
DB 107 RVLGSGIQL 117

RESULT 12
US-09-252-991A-32924
Sequence 32924, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 32924
LENGTH: 677
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32924

Query Match 42.7%; Score 41; DB 4; Length 677;
Best Local Similarity 47.4%; Pred. NO. 41;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKR 19
DB 90 KTLVDKILKSLTDRKE 108

RESULT 13
US-09-643-597-369
Sequence 369, Application US/09643597
Patent No. 6426072
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Fan, Liqun
APPLICANT: Kalos, Michael D.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: Hosken, Nancy
APPLICANT: Fanger, Gary R.
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aljun
APPLICANT: Skelky, Yasir A.W.
APPLICANT: Henderson, Robert A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.455C11
CURRENT APPLICATION NUMBER: US/09/643,597
CURRENT FILING DATE: 2000-08-21
NUMBER OF SEQ ID NOS: 369

SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 369
LENGTH: 708
TYPE: PRT
ORGANISM: Homo sapiens
US-09-643-597-369

Query Match 42.7%; Score 41; DB 4; Length 708;
Best Local Similarity 38.9%; Pred. No. 43;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 NLADRMGLSGVOEIRKQ 20
Db 373 NHOKRSIGTGIGIDLEK 390

RESULT 14
US-09-328-352-5192
Sequence 5192, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 5192
LENGTH: 295
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-5192

Query Match 42.2%; Score 40.5; DB 4; Length 295;
Best Local Similarity 47.6%; Pred. No. 18;
Matches 10; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

OY 3 NLADRM--LGLSGVOEIRKQ 20
Db 267 NLADKVALNTSISTIOEIAEQ 287

RESULT 15
US-09-634-238-247
Sequence 247, Application US/09634238
Patent No. 6544772
GENERAL INFORMATION:
APPLICANT: Glenn, Matthew
APPLICANT: Havukkala, Ilkka J.
APPLICANT: Bloksberg, Leonard, N.
APPLICANT: Lubbers, Mark W.
APPLICANT: Dekker, James
APPLICANT: Christenson, Anna C.
APPLICANT: Holland, Ross
APPLICANT: O'Toole, Paul W.
APPLICANT: Reid, Julian R.
APPLICANT: Coolbear, Timothy
TITLE OF INVENTION: Polynucleotides, materials incorporating
FILE REFERENCE: 11000, 1043U1
CURRENT APPLICATION NUMBER: US/09/634,238
CURRENT FILING DATE: 2000-08-08
NUMBER OF SEQ ID NOS: 422
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 247
LENGTH: 331
TYPE: PRT
ORGANISM: Lactobacillus rhamnosus
US-09-634-238-247

Query Match 41.7%; Score 40; DB 4; Length 331;
Best Local Similarity 42.1%; Pred. No. 26;

Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 2 INLADRMGLSGVOEIRKQ 20
Db 259 ISFLNQMONVDGYOEIRKQ 277

Search completed: October 1, 2003, 10:06:33
Job time: 13.7059 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	96	100.0	352	15	US-10-272-459-45
2	96	100.0	980	10	US-10-272-459-41
3	96	100.0	2227	10	US-09-929-955-12
4	96	100.0	2227	14	US-10-104-966-12
5	96	100.0	2227	14	US-10-135-988-2
6	96	100.0	2227	14	US-10-135-988-4
7	96	100.0	2227	14	US-10-135-988-6
8	45	46.9	632	15	US-10-156-761-13093
9	42	43.8	1250	10	US-09-801-368-364
10	41.5	43.2	1967	15	US-10-132-382-6
11	41.5	43.2	2013	15	US-10-132-382-2
12	41.5	43.2	2014	15	US-10-132-382-8
13	41.5	43.2	2040	15	US-10-132-382-4
14	41.5	43.2	366	10	US-09-870-759-72
15	41	42.7	366	12	US-09-751-708A-72

16	41	42.7	366	15	US-10-281-478-10	Sequence 10, Appl
17	41	42.7	605	9	US-09-815-242-12087	Sequence 12087, A
18	41	42.7	708	9	US-09-735-705-369	Sequence 369, App
19	41	42.7	708	10	US-09-850-716A-369	Sequence 369, App
20	41	42.7	708	10	US-09-897-778-369	Sequence 82, Appl
21	41	42.7	708	12	US-10-117-982-369	Sequence 369, App
22	41	42.7	1309	10	US-09-862-027-82	Sequence 82, Appl
23	40	41.7	140	10	US-09-738-626-4301	Sequence 4301, Ap
24	40	41.7	154	9	US-09-664-761-34931	Sequence 34931, A
25	40	41.7	316	12	US-09-907-218-76	Sequence 76, Appl
26	40	41.7	549	9	US-09-815-242-10680	Sequence 10680, A
27	40	41.7	615	10	US-09-738-626-6067	Sequence 6067, Ap
28	40	41.7	638	15	US-10-216-209-95	Sequence 95, Appl
29	40	41.7	643	9	US-09-826-660-25	Sequence 25, Appl
30	40	41.7	649	14	US-10-023-437-27	Sequence 27, Appl
31	40	41.7	1151	10	US-09-749-601A-10	Sequence 10, Appl
32	40	41.7	1151	11	US-09-912-697-33	Sequence 33, Appl
33	40	41.7	1186	9	US-09-826-660-23	Sequence 23, Appl
34	40	41.7	1207	11	US-09-988-462-7	Sequence 7, Appl
35	40	41.7	1770	10	US-09-801-368-298	Sequence 298, Appl
36	39	40.6	110	9	US-09-864-761-48195	Sequence 48195, A
37	39	40.6	270	12	US-10-237-496-88	Sequence 88, Appl
38	39	40.6	270	12	US-10-242-074-88	Sequence 88, Appl
39	39	40.6	270	12	US-10-242-505-88	Sequence 88, Appl
40	39	40.6	270	12	US-10-242-574-88	Sequence 88, Appl
41	39	40.6	270	12	US-10-243-261-88	Sequence 88, Appl
42	39	40.6	270	12	US-10-243-282-88	Sequence 88, Appl
43	39	40.6	270	12	US-10-243-402-88	Sequence 88, Appl
44	39	40.6	270	12	US-10-243-431-88	Sequence 88, Appl
45	39	40.6	270	12	US-10-243-164-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUNTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI/955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272.459
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kda
US-10-272-459-45

Query Match 100.0%; Score 96; DB 15; Length 352;

Best Local Similarity 100.0%; Pred. No. 2.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KINLADRMGLSGVQEIKEQ 20
|||||

Db 333 KINLADRMGLSGVQEIKEQ 352

RESULT 2
US-10-272-459-41

; Sequence 41, Application US/10272459
; Publication No. US20030124517A1

; GENERAL INFORMATION:
; APPLICANT: PICHUNTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

;; TITLE OF INVENTION: PROTEINS AND USES THEREOF
;; FILE REFERENCE: PRL7955.002 / 2301-17955
;; CURRENT APPLICATION NUMBER: US/10/272.459
;; CURRENT FILING DATE: 2002-10-15
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 41
;; LENGTH: 980
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
;; OTHER INFORMATION: of 115.5 kda
US-10-272-459-41

Query Match 100.0%; Score 96; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 7,4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVOEIKQ 20
IIIIIIIIIIIIIIIIIIII
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 3
US-09-929-955-12
;; Sequence 12, Application US/09929955
;; Patent No. US20020136740A1
;; GENERAL INFORMATION:
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPEP.23AUS2
;; CURRENT APPLICATION NUMBER: US/09/929,955
;; CURRENT FILING DATE: 2001-08-15
;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29
;; PRIOR APPLICATION NUMBER: 60/225,767
;; PRIOR FILING DATE: 2000-08-17
;; NUMBER OF SEQ ID NOS: 49
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 96; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVOEIKQ 20
IIIIIIIIIIIIIIIIIIII
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 4
US-10-104-966-12
;; Sequence 12, Application US/10104966
;; Publication No. US20020155124A1
;; GENERAL INFORMATION:
;; APPLICANT: Matti Salberg
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPEP.23AUSC1
;; CURRENT APPLICATION NUMBER: US/10/104,966
;; CURRENT FILING DATE: 2002-03-22

;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 96; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVOEIKQ 20
IIIIIIIIIIIIIIIIIIII
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 5
US-10-135-988-2
;; Sequence 2, Application US/10135988
;; Publication No. US20020176869A1
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US3
;; CURRENT APPLICATION NUMBER: US/10/135,988
;; CURRENT FILING DATE: 2002-04-29
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 96; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRLGLSGVOEIKQ 20
IIIIIIIIIIIIIIIIIIII
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 6
US-10-135-988-4
;; Sequence 4, Application US/10135988
;; Publication No. US20020176869A1
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US3
;; CURRENT APPLICATION NUMBER: US/10/135,988
;; CURRENT FILING DATE: 2002-04-29
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232


```

; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

```

```

Query Match          100.0%; Score 96; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1 KINLADRMGLSGVQEIKEQ 20
Db      961 KINLADRMGLSGVQEIKEQ 980

```

```

RESULT 7
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: PUNKHUSER, ANN M
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135, 988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

```

```

Query Match          100.0%; Score 96; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 KINLADRMGLSGVQEIKEQ 20
961 KINLADRMGLSGVQEIKEQ 980

```

```

RESULT 8
US-10-156-761-13093
; Sequence 13093, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109

```

```

; SEQ ID NO 13093
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13093

```

```

Query Match          46.9%; Score 45; DB 15; Length 622;
Best Local Similarity 52.6%; Pred. No. 27;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```

```

Oy      1 KINLADRMGLSGVQEIKE 19
Db      32 KSTLADRMGLTGVVDORQ 50

```

```

RESULT 9
US-09-801-368-364
; Sequence 364, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 364
; LENGTH: 1250
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-364

```

```

Query Match          43.8%; Score 42; DB 10; Length 1250;
Best Local Similarity 42.1%; Pred. No. 2e+02;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

```

```

Oy      2 INLADRMGLSGVQEIKEQ 20
Db      871 UNLDRITGFGVYINEIKR 889

```

```

RESULT 10
US-10-132-382-6
; Sequence 6, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1987
; TYPE: PRT

```

```
; ORGANISM: Homo sapiens
US-10-132-382-6

Query Match          43.2%; Score 41.5; DB 15; Length 1987;
Best Local Similarity 55.0%; Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 KINLADRMIG-LSGVQEIKE 19
   |||:| ||| |||
   |||:| ||| |||
DB 279 KLNLSDNHIGELPGVQSSDE 298

RESULT 11
US-10-132-382-2
; Sequence 2, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-2

Query Match          43.2%; Score 41.5; DB 15; Length 2013;
Best Local Similarity 55.0%; Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 KINLADRMIG-LSGVQEIKE 19
   |||:| ||| |||
   |||:| ||| |||
DB 305 KLNLSDNHIGELPGVQSSDE 324

RESULT 12
US-10-132-382-8
; Sequence 8, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2014
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-8

Query Match          43.2%; Score 41.5; DB 15; Length 2014;
Best Local Similarity 55.0%; Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 KINLADRMIG-LSGVQEIKE 19
   |||:| ||| |||
   |||:| ||| |||
DB 306 KLNLSDNHIGELPGVQSSDE 325

RESULT 13
US-10-132-382-4
; Sequence 4, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
```

```
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2040
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-4

Query Match          43.2%; Score 41.5; DB 15; Length 2040;
Best Local Similarity 55.0%; Pred. No. 4.4e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 1 KINLADRMIG-LSGVQEIKE 19
   |||:| ||| |||
   |||:| ||| |||
DB 332 KLNLSDNHIGELPGVQSSDE 351

RESULT 14
US-09-870-759-72
; Sequence 72, Application US/09870759
; Patent No. US2002017751A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-870-759-72

Query Match          42.7%; Score 41; DB 10; Length 366;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 7 RMLGSGVQEI 17
   |||:| ||| |||
   |||:| ||| |||
DB 107 RVLGISGLQEL 117

RESULT 15
US-09-751-708A-72
; Sequence 72, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-751-708A-72

Query Match          42.7%; Score 41; DB 12; Length 366;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
```

Fri Oct 2 09:14:00 2003

OY	7	RMIGSGVOEI	17
		1:11:11:11	
DB	107	RVIGISGLOEL	117

Search completed: October 1, 2003, 10:37:53
Job time : 25.1176 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 54.418 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-47
Perfect score: 129
Sequence: 1 ORLKYAQLSELSNEVLPPPKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq.19Jun03.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	18 AAW42930	Immunogenic Hepat
2	129	100.0	26	22 AAB69447	Synthetic HAV P2A
3	129	100.0	366	6 AAP50230	Sequence of hepat
4	129	100.0	854	6 AAP50287	Sequence encoded b
5	129	100.0	993	6 AAP50116	Sequence of Hepati
6	129	100.0	993	6 AAP50231	Sequence encoded b
7	129	100.0	1077	20 AAW95559	A partial hepatitis
8	129	100.0	1091	14 AAR32426	Translated from 5'
9	129	100.0	2227	11 AAR05697	Attenuated hepatitis

10	129	100.0	2227	18 AAW34074	Hepatitis A virus
11	129	100.0	2227	21 AAB18607	Amino acid sequenc
12	129	100.0	2227	21 AAB18609	Amino acid sequenc
13	129	100.0	2227	23 AAB31727	Wild-type Hepatit
14	129	100.0	2227	23 AAB31729	Attenuated Hepatit
15	129	100.0	2227	23 AAB19899	Hepatitis A virus
16	129	100.0	2227	24 ABO08639	Hepatitis A virus
17	129	100.0	2227	24 ABO08641	Wild type human he
18	124	96.1	2227	7 AAP60066	Attenuated hepati
19	124	96.1	2227	1 AAB18608	Sequence of viral
20	124	96.1	2227	23 AAB31728	Amino acid sequenc
21	124	96.1	2227	24 ABO08640	Hepatitis A virus
22	117	90.7	839	12 AAB15629	Attenuated (pass3
23	103	79.8	20	18 AAW42924	Capsid region of c
24	103	79.8	21	22 AAB69441	Immunogenic Hepat
25	65	50.4	20	18 AAW42925	Synthetic HAV P2A
26	65	50.4	21	22 AAB69442	Immunogenic Hepat
27	51	39.5	150	22 AAB68507	Synthetic HAV P2A
28	49	38.0	440	22 AAB63362	Drosophila melanog
29	49	38.0	444	18 AAW20624	Drosophila melanog
30	48	37.2	187	21 AAB32794	H. pylori cytoplas
31	48	37.2	221	22 AAB68504	Eucalyptus grandis
32	48	37.2	1176	22 AAB59592	Drosophila melanog
33	47.5	36.8	86	21 AAG03559	Drosophila melanog
34	47.5	36.8	530	21 AAB42529	Human secreted pro
35	47.5	36.8	530	23 AAO26501	Human ORF2293
36	47.5	36.8	530	23 AAB68940	Human glycoprotein
37	46.5	36.0	23	12 AAR15517	Human polyprotein
38	46.5	36.0	137	14 AAR53388	BMP-8 peptide. Bo
39	46.5	36.0	139	13 AAR27288	Murine osteogenic
40	46.5	36.0	139	14 AAR33401	Mature murine oste
41	46.5	36.0	139	14 AAR33924	Mouse mature OP-2.
42	46.5	36.0	139	15 AAR46735	Morphogen mop-2.
43	46.5	36.0	139	15 AAR60970	Mouse osteogenic p
44	46.5	36.0	139	17 AAW00224	Murine OP-2. Mus
45	46.5	36.0	139	18 AAW40183	Mouse mature osteo

ALIGNMENTS

RESULT 1	AAW42930	standard; peptide; 25 AA.
ID	AAW42930	
XX	AAW42930:	
AC	28-APR-1998	(first entry)
XX		
DE	Immunogenic Hepatitis A virus peptide YK-1665.	
XX		
KW	Immunogenic peptide; immunogenic epitope; P2A protein;	
XX	Immune response; antibody.	
OS	Synthetic.	
XX	Hepatitis A virus.	
PN	W09740147-A1.	
XX		
PD	30-OCT-1997.	
XX		
PF	18-APR-1997; 97MO-US06891.	
XX		
PR	19-APR-1996; 96US-0015644.	
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
PI	Fields HA, Khudayakov YE;	
XX		
DR	WPI; 1997-535831/49.	
XX		
PT	Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an	
	Immune response to HAV in a mammal or to detect the presence of	

PT antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. Compositions
CC containing the peptides can be used to induce an immune response to HAV
CC in a mammal. The peptides can also be used to detect the presence of
CC antibodies against HAV in mammalian serum. The peptides can also be used
CC to make an antibody against HAV by administering the peptide to a
CC mammal.
XX
SQ Sequence 25 AA;
XX
Query Match 100.0%; Score 129; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 8.3e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 QRLRYAOEELSNEVLPPEPRKKMGLE 25
1 QRLRYAOEELSNEVLPPEPRKKMGLE 25
RESULT 2
AAB69447
ID AAB69447 standard; Peptide: 26 AA.
XX
AC AAB69447;
XX
DT 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 47.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
PD 25-JAN-2001.
XX
PE 14-JUL-2000; 2000WO-US19267.
XX
PR 15-JUL-1999; 99US-0144412.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Fields HA, Khudyakov YE;
XX
DR WPI: 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 98; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
SQ Sequence 26 AA;
XX
Query Match 100.0%; Score 129; DB 22; Length 26;
Best Local Similarity 100.0%; Pred. No. 8.7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 QRLRYAOEELSNEVLPPEPRKKMGLE 25
1 QRLRYAOEELSNEVLPPEPRKKMGLE 25
RESULT 3
AAP50230
ID AAP50230 standard; Protein: 366 AA.
XX
AC AAP50230;
XX
DT 28-NOV-1991 (first entry)
XX
DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
OS Hepatitis A virus.
XX
PN EP138704-A.
XX
PD 24-APR-1985.
XX
PE 09-OCT-1984; 84EP-0402025.
XX
PR 02-MAR-1984; 84US-0585942.
XX
PR 14-OCT-1983; 83US-0541836.
XX
PA (MERI) MERCK & CO INC.
XX
PI Hughes JV, Scolnick EM, Tomassini JE;
XX
DR WPI: 1985-100818/17.
XX
DR N-PSDB; AAN50274.
XX
PT New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus
XX
PS Claim 21; Page 46-48; 49pp; English.
XX
CC VPI is isolated by solubilisation of the intact virus in an aq.
CC anionic surfactant and a reducing agent. The viral proteins are sepd.
CC and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ Sequence 366 AA;
XX
Query Match 100.0%; Score 129; DB 6; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 QRLRYAOEELSNEVLPPEPRKKMGLE 25
1 QRLRYAOEELSNEVLPPEPRKKMGLE 343
RESULT 4
AAP50287
ID AAP50287 standard; Protein: 854 AA.
XX
AC AAP50287;
XX
DT 25-MAR-2003 (updated)

```

DT      30-NOV-1991 (first entry)
XX
XX      Sequence encoded by hepatitis A virus (HAV) cDNA from near the
DE genome 5' terminus to the end of the area corresponding to the
DE capsid protein region of poliovirus RNA.
XX
XX      Hepatitis A virus assay; antigen; antibody.
XX
XX      Hepatitis A virus.
OS
XX      MO8501517-A.
XX
XX      11-APR-1985.
XX
XX      27-SEP-1984; 84MO-US01552.
XX
XX      30-SEP-1983; 83US-0537911.
XX
XX      (MMSI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX      Titechurst JR, Baltimore D, Feinstein SM, Purcell RH;
XX      Racanello VR;
XX      WPI; 1985-098846/16.
XX
XX      N-PSDB; AAN50330.
XX
XX      New hepatitis A virus cDNA - useful in assays for the virus and
XX      for prodn. of the viral antigen and antibodies to it
XX
XX      Example; Fig 7; 60pp: English.
XX
XX      The inventors claim HAV cDNA and a method for producing it, whereby
XX      large amts. can be obt'd. economically. The cDNA is useful in the
XX      assay for detection of HAV quickly and easily and with high
XX      sensitivity and specificity. The HAV cDNA is also used in the prodn.
XX      of HAV antigen or antibodies to it. The antibodies may be monoclonal.
XX      (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence      854 AA:
SQ
XX
XX      Query Match      100.0%; Score 129; DB 6; Length 854;
XX      Best Local Similarity 100.0%; Pred. No. 3.8e-11;
XX      Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
XX      1 ORLKYAOEELSNEVLP PPRKKMGLE 25
XX      |||||||
XX      810 ORLKYAOEELSNEVLP PPRKKMGLE 834
XX
XX      5
XX      AAF50116
XX      AAF50116 standard; Protein; 993 AA.
XX
XX      AAF50116;
XX
XX      25-MAR-2003 (updated)
XX      30-SEP-1991 (first entry)
XX
XX      Sequence of Hepatitis A virus (HAV) immunogenic peptides
DE VP-1, VP-2, VP-3 and VP-4.
XX
XX      Antigenic protein; immunogen; vaccine.
XX
XX      Hepatitis A virus (strain CR326).
XX
XX      EPI54587-A.
XX
XX      11-SEP-1985.
XX
XX      27-FEB-1985; 85EP-0400369.
XX
XX      02-MAR-1984; 84US-0585818.
XX

```

```

PA      (MERI ) MERCK & CO INC.
XX
PI      Linemeyer DL, Menke JG, Rueben RG, Mitra SW;
XX
DR      WP1; 1985-224964/37.
XX
DR      N-PSDB; AAN50139.
XX
PT      New nucleotide sequences coding for hepatitis A virus antigens -
PT      useful for eliciting normal immune response and in vaccines for
PT      protecting against the virus
XX
PS      Example; Page 11-17; 32pp; English.
XX
CC      Within the sequence in AAN50139 is encoded the information necessary
CC      to make the antigenic proteins of HAV. The sequences encoding for
CC      the structural proteins begin at base 403. The key sub-unit
CC      sequences within VP-1, designated Sequences I, II, III, IV, and V,
CC      start, respectively at 1882, 1963, 1999, 2146, 2347. Other
CC      nucleotide sequences which are valuable as encoding antigenic
CC      proteins are the sequences from base 1749 to base 2722; from base
CC      1487 to base 2980 and from base 1644 to base 2722. The sequence from
CC      base 1749 to base 2722 is esp. valuable as a vector for producing
CC      antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
CC      translation of a stop codon.
CC      (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ      Sequence    993 AA;
XX
Query Match          100.0%; Score 129; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 4,4e-11;
Matches   25; Conservative    0; Mismatches    0; Indels    0; Gaps    0.

OY      1 QRLKYAOEELSENVLP PPRKMGLEF 25
        |||||||
DB       946 QRLKYAOEELSENVLP PPRKMGLEF 970

RESULT 6
AAP50231
ID      AAP50231 standard; Protein; 993 AA.
XX
AC      AAP50231;
XX
DT      28-NOV-1991 (first entry)
XX
DE      Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE      including surface protein (VP-1).
XX
KM      Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KM      diagnostic assay.
XX
OS      Hepatitis A virus.
XX
FH      Key              Location/Qualifiers
FT      Protein         628..993
FT                                     /note= "claimed; x denotes translated stop codons
FT                                     and unspecified triplets"
PN      EPI38704-A.
XX
PD      24-APR-1985.
XX
PF      09-OCT-1984;    84EP-0402025.
XX
PR      02-MAR-1984;    84US-0585942.
PR      14-OCT-1983;    83US-0541836.
XX
PA      (MERI ) MERCK & CO INC.
XX
PI      Hughes JV, Scolnick EM, Tomassini JE;
XX
DR      WPI. 1985-100818/17.
DR      N-PSDB; AAN50274.

```

```

XX  New hepatitis A virus surface protein - useful for binding to
PT  neutralising antibodies to the virus
XX
XX  Disclosure: Page 17-23; 49pp; English.
PS
XX  WPI is isolated by solubilisation of the intact virus in an aq.
CC  anionic surfactant and a reducing agent. The viral proteins are sepd.
CC  and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ  Sequence 993 AA:

Query Match 100.0%; Score 129; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 4,4e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 QRLKYAOEELSNEVLPPEPRKMGLE 25
|||||
Db 946 QRLKYAOEELSNEVLPPEPRKMGLE 970

ID AAW95559 standard; Protein; 1077 AA.
XX
AC AAW95559;
XX
DT 28-APR-1999 (first entry)
XX
DE A partial hepatitis A virus (HAV) protein.
XX
KM Hepatitis A virus protein; HAV; P2 region;
KW cell-culture-adapted HAV strain; infection; accelerated growth.
XX
OS Hepatitis A virus.
XX
PN US5849562-A.
XX
PD 15-DEC-1998.
XX
PF 06-JUN-1995; 95US-0468926.
XX
PR 06-NOV-1991; 91US-0788262.
PR 30-SEP-1983; 83US-0537911.
PR 27-SEP-1984; 84US-0654942.
PR 06-OCT-1988; 88US-0256135.
PR 06-JUN-1995; 95US-0468926.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH;
XX
XX WPI; 1999-094412/08.
XX
XX N-PSDB; AAX01006.
XX
XX Chimeric hepatitis A virus strains - with P2 region from
PT cell-culture-adapted strain in wild-type genome
XX
XX Disclosure: Fig 7A-L; 36pp; English.
XX
XX The present sequence represents a partial hepatitis A virus (HAV)
CC protein. The specification describes a DNA construct consisting
CC of a wild-type HAV genome in which the P2 region is replaced by the
CC P2 region from a cell-culture-adapted HAV strain. The construct is
CC used to demonstrate that mutations in the P2 region of a
CC cell-culture-adapted HAV strain are sufficient for establishment of
CC infection and accelerated growth in cell culture.
XX
SQ Sequence 1077 AA:

Query Match 100.0%; Score 129; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 4,8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 QRLKYAOEELSNEVLPPEPRKMGLE 25
DB 1033 QRLKYAOEELSNEVLPPEPRKMGLE 1057

RESULT 8
AAR32426
ID AAR32426 standard; Protein; 1091 AA.
XX
AC AAR32426;
XX
DT 25-MAR-2003 (updated)
DT 17-DEC-2001 (updated)
DT 10-JUN-1993 (first entry)
XX
DE Translated from 5' region of Hepatitis A virus genomic clone.
XX
XX HAV HM-175; chronic liver disease; picornavirus.
XX
XX Hepatitis A virus.
XX
XX OS
XX FH Key Location/Qualifiers
XX FT Region 238..1091
XX FT /label= ORF
XX FT /note= "second putative initiation codon at
XX FT position 240"
XX FT Region 1..711
XX FT /note= "X's correspond to nonsense codons,
XX FT 1.e. this region is not an ORF"
XX
XX USN7/88262-N.
XX
XX PD 15-DEC-1992.
XX
XX PF 30-SEP-1983; 83US-0536911.
XX
XX PR 27-SEP-1984; 84US-0654942.
XX PR 06-OCT-1988; 88US-0256135.
XX PR 30-SEP-1983; 83US-0536911.
XX PR 06-NOV-1991; 91US-0788262.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX PA Ricehurst JR, Baltimore D, Feinstein SM, Purcell RH, Racanietello VR;
XX PI Baroudy BM, Emerson SU;
XX DR WPI; 1993-067429/08.
XX
XX N-PSDB; AAO36934.
XX
XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
XX of antigen and antibodies
XX
XX Disclosure: Fig 7; 65pp; English.
XX
XX HAV virion RNA was extracted from the livers of marmosets which had
CC been inoculated with HAV (the HAV had previously been passaged twice
CC in marmosets). The RNA was used to prepare ds cDNA clones by
CC standard methods. Clones contg. inserts which hybridised to RNA from
CC HAV-infected African Green Monkey kidney cells were selected for
CC further analysis. A 7.4kb restriction map (about 9% of the HAV
CC genome) was constructed from 5 overlapping inserts. The sequence of
CC the first 3.3kb (approx.) from the 5'-terminus was determined. An
CC amino acid sequence was deduced from the entire clone and an open
CC reading frame was identified starting at position 238. A comparison
CC of the predicted HAV amino acid sequences with the known capsid
CC protein sequences of other picornaviruses (poliovirus, foot and
CC mouth disease virus and encephalomyelitis virus) revealed areas of
CC local homology.
CC (Note: Revised entry submitted to correct the patent number format of
CC US government-owned NIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/nlis-us.html.)

```


CC (Updated on 25-MAR-2003 to correct PI field.)
XX
SQ Sequence 1091 AA;
Query Match 100.0%; Score 129; DB 14;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25
DB 1047 ORLKYAOEELSNEVLPPEPRKMKGLF 1071
RESULT 9
AAR05697
ID AAR05697 standard; protein; 2227 AA.
XX
AC AAR05697;
XX
XX 25-MAR-2003 (updated)
XX 15-AUG-1990 (first entry)
DE Attenuated hepatitis A virus.
XX
XX Hepatitis A virus; vaccine; attenuated.
XX
OS Hepatitis A virus, strain HM-175.
XX
XX
FH Key Location/Qualifiers
FT 1..23
FT /label-VP4 = 1A
FT 24..245
FT /label-VP2 = 1B
FT 246..491
FT /label-VP3 = 1C
FT 492..791
FT /label-VP1 = 1D
FT 792..980
FT /label-2A
FT 981..1087
FT /label-2B
FT 1088..1422
FT /label-2C
FT 1423..1496
FT /label-3A
FT 1497..1519
FT /label-3B = VPg
FT 1520..1738
FT /label-3C
FT 1739..2227
FT /label-3D
XX
XX US4894228-A.
XX
XX
XX 16-JAN-1990.
XX
XX
XX 12-JUL-1988; 88US-0217824.
XX
XX 12-JUL-1988; 88US-0217824.
XX 12-JUL-1988; 88US-0652967.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
XX Daemer RJ, Gust ID;
XX
XX WPI: 1990-075557/10.
XX N-PSDB: AAG03512.
XX
XX Vaccine against hepatitis A virus infection comprises novel
XX attenuated hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 18pp; English.
XX
XX

XX
XX The attenuated HAV is useful for inducing protective immunity against
XX HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
XX several nucleotide changes distributed throughout the genome, is
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is
XX suitable for use as an HAV vaccine. It is noted that not all the changes
XX are necessary for attenuation and use as a vaccine.
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25
DB 810 ORLKYAOEELSNEVLPPEPRKMKGLF 834
RESULT 10
AAW34074
ID AAW34074 standard; Protein; 2227 AA.
XX
XX AAW34074;
XX
XX 27-APR-1998 (first entry)
XX
XX
DE Hepatitis A virus HM-175 protein sequence.
XX
XX
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
XX infection; vaccine.
XX
XX
OS Hepatitis A virus HM-175.
XX
XX
FH Key Location/Qualifiers
FT 1..23
FT /label-VP4
FT 24..245
FT /label-VP2
FT 246..491
FT /label-VP3
FT 492..791
FT /label-VP1
FT 792..980
FT /label-2A
FT 981..1087
FT /label-2B
FT 1088..1422
FT /label-2C
FT 1423..1496
FT /label-3A
FT 1497..1519
FT /label-3B
FT 1520..1738
FT /label-3C
FT 1739..2227
FT /label-3D
XX
XX W09740166-A2.
XX
XX
XX 30-OCT-1997.
XX
XX
XX 18-APR-1997; 97WO-US06506.
XX
XX 19-APR-1996; 96US-0015642.
XX
XX
XX (USSH) US SEC DEPT HEALTH.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH, Raychaudhuri G;
XX
XX

DR	WPI; 1997-535850/49.
DR	N-PSDB; AAT93023.
XX	
PT	Human attenuated HAV genome containing simian HAV 2C gene - useful as vaccines against HAV infection
XX	
PS	Disclosure; Fig 13A-D; 66pp; English.
CC	This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (I) comprises a genome of HAV, where the 2C gene has been replaced by a corresponding region of the 2C gene of a human attenuated HAV genome in which a from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (I) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal.
CC	
CC	
SO	Sequence 2227 AA;
QY	Query Match 100.0%; Score 129; DB 18; Length 2227; Best Local Similarity 100.0%; Pred. NO. 1.1e-10; Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 ORLKYAOEELSNEVLPPPRKKKGIF 25 810 ORLKYAOEELSNEVLPPPRKKKGIF 834
RESULT 11	
ID	AAB18607
AC	AAB18607 standard; Protein; 2227 AA.
XX	
DT	15-JAN-2001 (first entry)
XX	
DE	Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX	
KX	HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX	
XX	Hepatitis A virus.
XX	
PD	US6113912-A.
XX	
XX	05-SEP-2000.
XX	
XX	07-JUN-1995; 95US-0475886.
XX	
PR	18-SEP-1992; 92US-0947338.
PR	17-SEP-1993; 93WO-US08610.
PR	10-MAR-1995; 95US-0397232.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
PI	D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX	
DR	WPI; 2000-586464/55.
DR	N-PSDB; AAA75476.
XX	
XX	Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type -
XX	
XX	Disclosure; Fig 6A-K; 72pp; English.

```

CC CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
CC
XX
XX
S0 Sequence 2227 AA;
    Query Match 100.0%; Score 129; DB 21; Length 2227;
    Best Local Similarity 100.0%; Pred. No. 1.1e-10;
    Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAOEELSNEVLPPRRKKGLF 25
    |||||||||||||||||||||||||
DB 810 QRLKYAOEELSNEVLPPRRKKGLF 834
    |||||||||||||||||||||||||

RESULT 12
AAB18609
ID AAB18609 standard; Protein; 2227 AA.
XX
XX AAB18609;
AC
XX 15-JAN-2001 (first entry)
DT
XX
XX Amino acid sequence of live attenuated Hepatitis A virus 4380.
DE
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX HAV 4380.
XX
XX Hepatitis A virus.
OS
XX US6113912-A.
PN
XX
XX 05-SEP-2000.
PD
XX
XX 07-JUN-1995; 95US-0475886.
PE
XX
XX 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93MO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI: 2000-586464/55.
XX N-PSDB; AAA75478.
DR
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type -
XX
XX Disclosure: Columns 93-104: 72pp; English.
XX
XX The present sequence is derived from a live attenuated hepatitis A
XX virus (HAV) of the invention, designated HAV 4380. The sequence is
XX produced by modifying wild type HAV strain HM-174. The HAV of the
XX invention are adapted to growth in the human fibroblast-like cell
XX line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
XX appropriate attenuation. It is useful as a live vaccine for prophylaxis
XX of hepatitis A in humans and other primates.
XX
XX
S0 Sequence 2227 AA;
    Query Match 100.0%; Score 129; DB 21; Length 2227;
    Best Local Similarity 100.0%; Pred. No. 1.1e-10;
    Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAOEELSNEVLPPRRKKGLF 25
    |||||||||||||||||||||||||
DB 810 QRLKYAOEELSNEVLPPRRKKGLF 834
    |||||||||||||||||||||||||

```

RESULT 13
 ABG31727
 ID ABG31727 standard; Protein: 2227 AA.
 XX
 AC ABG31727;
 XX
 DT 29-NOV-2002 (first entry)
 XX
 DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
 XX
 KM Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
 XX
 OS Hepatitis A virus strain HM-175.
 XX
 PN US6423318-B1.
 PD 23-JUL-2002.
 31-AUG-2000; 2000US-0653499.
 07-JUN-1995; 95US-0475886.
 PR 17-SEP-1993; 93US-0397232.
 PR 17-SEP-1993; 93WO-US08610.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Funhouser AM, Emerson SU, Purcell RH, D'Hondt E;
 DR WPI; 2002-680946/73.
 DR N-PSDB; ABS52787.
 XX
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 in MRC-5 cells, useful for preparing a vaccine against HAV infection -
 XX
 PS Disclosure; Fig 6; 71pp; English.
 XX
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC strain HM-175 polypeptide.
 CC
 SQ Sequence 2227 AA:
 Query Match 100.0%; Score 129; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ORLKYAOEELSNVLP...PPRKMGLF 25
 810 ORLKYAOEELSNVLP...PPRKMGLF 834
 RESULT 14
 ABG31729
 ID ABG31729 standard; Protein: 2227 AA.
 XX
 AC ABG31729;
 XX
 DT 29-NOV-2002 (first entry)
 XX
 DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
 XX
 KM Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 KM HAV 4380.
 XX
 OS Hepatitis A virus strain HM-175.
 XX
 PN US6423318-B1.

PD 23-JUL-2002.
 XX
 PF 31-AUG-2000; 2000US-0653499.
 XX
 PR 07-JUN-1995; 95US-0475886.
 PR 17-SEP-1993; 93US-0397232.
 PR 17-SEP-1993; 93WO-US08610.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Funhouser AM, Emerson SU, Purcell RH, D'Hondt E;
 DR WPI; 2002-680946/73.
 DR N-PSDB; ABS52789.
 XX
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 in MRC-5 cells, useful for preparing a vaccine against HAV infection -
 XX
 PS Disclosure; Column 93-104; 71pp; English.
 XX
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents an attenuated
 CC hepatitis A virus 4830 polypeptide.
 CC
 SQ Sequence 2227 AA:
 Query Match 100.0%; Score 129; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ORLKYAOEELSNVLP...PPRKMGLF 25
 810 ORLKYAOEELSNVLP...PPRKMGLF 834
 RESULT 15
 AAE19899
 ID AAE19899 standard; Protein: 2227 AA.
 XX
 AC AAE19899;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Hepatitis A virus (HAV) protein.
 XX
 KM Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
 KM cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
 XX
 OS Hepatitis A virus.
 XX
 PN WO200213855-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 15-AUG-2001; 2001WO-1B01808.
 XX
 PR 17-AUG-2000; 2000US-225767P.
 PR 29-AUG-2000; 2000US-229175P.
 PR 03-NOV-2000; 2000US-0705547.
 XX
 PA (TRIP-) TRIPEP AB.
 XX
 PI Sallberg M, Hultgren C;
 DR WPI; 2002-241837/29.
 DR N-PSDB; AAD31766.
 XX
 PT Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus

XX Claim 11; Page 82-87; 120pp; English.
 PS
 CC The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein.
 XX

Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 23; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.1e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOELSNVLPPIPRKMKGLF 25
 |||||
 810 ORLKYAOELSNVLPPIPRKMKGLF 834

Search completed: October 1, 2003, 09:56:46
 Job time : 55.4118 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 16.0294 seconds
(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129

Sequence: 1 QRLRYAQBELSNVLPPEPRKMKGLF 25

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

al number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pIR.76:*
2: pIR1:*
3: pIR2:*
4: pIR3:*
5: pIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	2 S04137	genome polyprotein
2	129	100.0	852	2 GNNYHA	genome polyprotein
3	129	100.0	1358	2 A03905	genome polyprotein
4	129	100.0	2227	1 GNNYHM	genome polyprotein
5	129	100.0	2227	1 GNNYHM	genome polyprotein
6	124	96.1	2227	1 GNNYHM	genome polyprotein
7	123	95.3	2227	1 GNNYHM	genome polyprotein
8	119	92.2	2230	1 GNNYSA	genome polyprotein
9	117	90.7	839	1 GNNYSA	genome polyprotein
10	56.5	43.8	346	2 S74448	genome polyprotein
11	52	40.3	859	2 T43701	regulatory protein
12	52	40.3	1193	2 E88445	DNA-directed RNA P
13	51	39.5	1119	2 T50995	protein C26E6.4 (1
14	51	39.5	6642	2 T29757	related to cytoske
15	50.5	39.1	443	2 E82046	protein UNC-89 - C
16	49.5	38.4	159	2 C72210	proteinase HslVU
17	49.5	38.4	554	2 A56730	conserved hypochet
18	49	38.0	55	2 P00433	carl protein - Pod
19	49	38.0	55	2 P00434	genome polyprotein
20	49	38.0	56	2 P00428	genome polyprotein
21	49	38.0	56	2 P00427	genome polyprotein
22	49	38.0	56	2 P00429	genome polyprotein
23	49	38.0	56	2 P00432	genome polyprotein
24	49	38.0	56	2 P00430	genome polyprotein
25	49	38.0	442	2 A71969	probable histidine
26	49	38.0	465	2 H86482	protein F5U5.11 (1
27	49	38.0	1174	2 S28976	DNA-directed RNA P
28	48.5	37.6	2278	1 S56274	FAB1 protein - yea
29	48	37.2	523	2 D8538	hypothetical prote

30	48	37.2	523	2 H90687	hypothetical prote
31	48	37.2	1176	2 A27826	DNA-directed RNA P
32	48	37.2	1191	2 S65068	DNA-directed RNA P
33	48	37.2	1210	2 S35548	DNA-directed RNA P
34	47.5	36.8	285	2 T15133	hypothetical prote
35	47.5	36.8	1626	2 T09271	probable tail comp
36	47	36.4	223	2 P00514	hemagglutinin HA2
37	47	36.4	254	2 F75575	hypothetical prote
38	47	36.4	578	2 S03299	hemagglutinin prec
39	47	36.4	592	1 LIRY	actin-binding prot
40	46.5	36.0	361	2 G82530	hypothetical prote
41	46.5	36.0	486	2 T10089	clitrin - sweet ora
42	46.5	36.0	634	2 E86293	T24D18.1 protein -
43	46	35.7	163	2 PC4186	hemagglutinin 2 ch
44	46	35.7	262	2 A70428	hypothetical prote
45	46	35.7	338	2 I56893	transcription fact

ALIGNMENTS

RESULT 1

S04137 genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)

C:Species: human hepatitis A virus

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000

C:Accession: S04137

R:Anderson, A.P.; Lau, P.C.K.; Chaudhary, R.

Nucleic Acids Res. 17, 3594, 1989

A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir

A:Reference number: S04137; MID:89263805; PMID:2542903

A:Accession: S04137

A:Molecule type: mRNA

A:Residues: 1-341 <AND>

A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576

C:Genetics:

A:Gene: VP1

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; polyprotein

F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match	100.0%	Score 129;	DB 2;	Length 341;
Best Local Similarity	100.0%	Pred. No. 9.4e-12;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
DB	314	QRLRYAQBELSNVLPPEPRKMKGLF	338	

RESULT 2

GNNYHA genome polyprotein - human hepatitis A virus (strain CR326) (fragment)

N:contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C:Accession: A03904

R:Lindegger, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra,

J. Virol. 54, 247-255, 1985

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 129; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 2,7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25
|||||

Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 3
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein
C;Species: human hepatitis A virus
C:Date: 31-Mar-1991 #sequence,revision 31-Mar-1991 #text_change 15-Nov-1996
A:Accession: A03905
R:Baroudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Maisel Jr., J.V.; Purcell, R.H.; Feinstor
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:Reference number: A03905; MUID:85166289; PMID:2984684
A:Accession: A03905
A:Molecule type: genomic RNA
A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 129; DB 2; Length 1338;
Best Local Similarity 100.0%; Pred. No. 4,6e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25
|||||

Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 4
GNNYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
B; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence,revision 31-Mar-1988 #text_change 16-Jul-1999
A:Accession: A25981
R:Conen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8e-11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25
|||||

Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 5
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48); protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence,revision 30-Jun-1987 #text_change 16-Jul-1999
A:Accession: A03903
R:Marjarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25
|||||

Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 6
GNNYMK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core
NA polymerase (EC 2.7.7.48); protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence,revision 30-Jun-1988 #text_change 16-Jul-1999
A:Accession: A94149; A25914; A84508
R:Conen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstone, S.M.; Purcell
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8e-11;

F:1423-1484/Product: protein 3A #status predicted <P3A>
 F:1485-1507/Product: protein 3B #status predicted <P3B>
 F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
 F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 96.1%; Score 124; DB 1; Length 2227;
 Best Local Similarity 96.0%; Pred. No. 4.6e-10;
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25
 |||||
 DB 810 ORLKYAOEELSNEVLPPPKMKGLF 834

RESULT 7

GNNYB

genome polypeptide - human hepatitis A virus (strain MB)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; protein 1F; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D

Species: human hepatitis A virus
 Note: host Homo sapiens (man)

Accession: J50303
 Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996

R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhardt, Virus Res. 8, 153-171, 1987

A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat

A:Reference number: J50303; MUID:88045071; PMID:2823500

A:Accession: J50303
 A:Molecule type: genomic RNA

A:Residues: 1-2227 <PAU>

A:Cross-references: EMBL:M20723

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hyd

F:1-23/Product: coat protein 1A #status predicted <P4>

F:24-246/Product: coat protein 1B #status predicted <VP2>

F:492-836/Product: coat protein 1C #status predicted <VP3>

F:837-980/Product: coat protein 1D #status predicted <VP1>

F:981-1108/Product: core protein 2A #status predicted <P2A>

F:1109-1438/Product: core protein 2B #status predicted <P2B>

F:1439-1496/Product: core protein 2C #status predicted <P2C>

F:1497-1519/Product: genome-linked protein 3A #status predicted <P3A>

F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>

F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 95.3%; Score 123; DB 1; Length 2227;
 Best Local Similarity 96.0%; Pred. No. 6.5e-10;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25
 |||||
 DB 810 ORLKYAOEELSNEVLPPPKMKGLF 834

RESULT 8

GNNYA

genome polypeptide - simian hepatitis A virus (strain AGM-27)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro

C:Species: simian hepatitis A virus

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000

C:Accession: A50470; S04885; S03965

R:Tsarev, S.A.
 submitted to JIPID, April 1991

A:Reference number: A50470

A:Accession: A50470

A:Molecule type: genomic RNA

A:Residues: 1-2230 <TSA>

A:Cross-references: GB:D00924; NID:9222597; PIDN:BA00766.1; PID:9222598

J:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.

A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an

A:Reference number: J01080; MUID:91311420; PMID:1649901

A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chlzh
 submitted to the EMBL Data Library, May 1989

A:Reference number: S04885
 A:Accession: S04885

A:Molecule type: genomic RNA

A:Residues: 1750-2164 <BAL1>

A:Cross-references: EMBL:X15461; NID:961971; PIDN:CAA3490.1; PID:9930268

R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chlzh

FEBS Lett. 247, 425-428, 1989

A:Title: Variations in genome fragments coding for RNA polymerase in human and simian

A:Reference number: S03965; MUID:89232168; PMID:2541023

A:Accession: S03965

A:Molecule type: genomic RNA

A:Residues: 1960-2164 <BAL2>

A:Cross-references: EMBL:X15461

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; polypeptide

F:1-27/Product: coat protein 1A #status predicted <S1A>

F:28-249/Product: coat protein 1B #status predicted <S1B>

F:250-495/Product: coat protein 1C #status predicted <S1C>

F:496-795/Product: coat protein 1D #status predicted <S1D>

F:796-984/Product: core protein 2A #status predicted <S2A>

F:985-1091/Product: core protein 2B #status predicted <S2B>

F:1092-1426/Product: core protein 2C #status predicted <S2C>

F:1427-1498/Product: protein 3A #status predicted <P3A>

F:1499-1521/Product: protein 3B #status predicted <P3B>

F:1522-1741/Product: protein 3C #status predicted <P3C>

F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 92.2%; Score 119; DB 1; Length 2230;
 Best Local Similarity 88.0%; Pred. No. 2.6e-09;

Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25
 |||||
 DB 814 ORLKYAOEELSNEVLPPPKMKGLF 838

RESULT 9

GNNY2

genome polypeptide - simian hepatitis A virus (strain CY-145) (fragment)

N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat prot

C:Species: simian hepatitis A virus

A:Note: host Macaca fascicularis (cynomolgus macaque)

C:Accession: J01180

R:Malman, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.

J. Gen. Virol. 72, 1685-1689, 1991

A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus

A:Reference number: J01180; MUID:91311421; PMID:1649902

A:Accession: J01180

A:Molecule type: genomic RNA

A:Residues: 1-839 <NA1>

A:Cross-references: GB:M59286; NID:9329599; PIDN:AAA5473.1; PID:9555083

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; glycoprotein; polypeptide

F:1-23/Product: coat protein 1A #status predicted <VP0>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP1>

F:492-839/Product: coat protein 2A (fragment) #status predicted <P2>

F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.7%; Score 117; DB 1; Length 839;
 Best Local Similarity 84.0%; Pred. No. 1.7e-09;

Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25
 |||||
 DB 809 ORLKYAOEELSNEVLPPPKMKGLF 833

```

574448 regulatory protein pcr - Synechocystis sp. (strain PCC 6803)
N.Alternate names: protein sl11408
C.Species: Synechocystis sp.
A.Variety: PCC 6803
C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
Accession: S74448
R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-116, 1996
A.Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A.Reference number: S74322; MUID:97061201; PMID:8905231
A.Accession: S74448
A.Status: nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-346 <KAN>
A.Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL6600.1; PID:g165165
Genetics:
A.Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
Gene: pcr
Superfamily: Synechocystis regulatory protein pcr
Keywords: transcription regulation

Query Match 43.8%; Score 56.5; DB 2; Length 346;
Best Local Similarity 40.0%; Pred. No. 0.95;
Matches 14; Conservative 4; Mismatches 4; Indels 13; Gaps 1;

OY 1 QRLKAOEELSENYLPP-----RKMK 22
|||:|||||:|:|:|
DB 241 ERVKTAQETLVKQVLDPPSLAOLSRQVSLNERKLK 275

RESULT 11
T43701
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain [Imported] - Caenorhabditis el
C.Species: Caenorhabditis elegans
C.Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
Accession: T43701
R.Sidow, A.; Thomas, W.K.
Curr. Biol. 4, 596-603, 1994
A.Title: A molecular evolutionary framework for eukaryotic model organisms.
A.Reference number: Z22636; MUID:95041334; PMID:7955533
A.Accession: T43701
A>Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-859 <SID>
A.Cross-references: EMBL:U10333; NID:g520506; PIDN:AAA50224.1; PID:g520507
Superfamily: DNA-directed RNA polymerase 132k polypeptide
Keywords: nucleotidyltransferase

Query Match 40.3%; Score 52; DB 2; Length 859;
Best Local Similarity 62.5%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 QRLKAOEELSENYLP 16
|||:|||||:|:|:|
DB 141 QRITAREILLKEPLL 156

RESULT 12
E88445
protein C26E6.4 [Imported] - Caenorhabditis elegans
C.Species: Caenorhabditis elegans
C.Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-May-2001
Accession: E88445
R.anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A.Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A.Reference number: A75000; MUID:99065613; PMID:9851916
A.Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A.Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; ano
A.Accession: E88445
```

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 11193 <STO>
A:Cross-references: GB:chr_III; PIDN:AAA21158.1; PID:q532805; GSPDB:GN00021; CESP:C26
C:Genetics:
A:Gene: C26B6.4
A:Map position: 3
C:Superfamily: DNA-directed RNA polymerase 132k polypeptide

Query Match 40.3%; Score 52; DB 2; Length 1193;
Best Local Similarity 62.5%; Pred. NO. 19;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORLKYAOELSENVLP 16
||:||||:| |:||
Db 340 QRIKARELLQKELLP 355

RESULT 13
T50995
related to cytoskeleton assembly control protein SLA1 [Imported] - Neurospora crassa
N:Alternate names: protein B7F18.140
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T50995
R:Schulte, U.; Aism, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, July 2000
A:Reference number: 225286
A:Accession: T50995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 11119 <SCH>
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.140
A:Experimental source: BAC clone B7F18; strain OR74A.
C:Genetics:
A:Gene: NCSP:B7F18.140
A:Map position: 6
A:Introns: 66/3; 123/2; 495/1

Query Match 39.5%; Score 51; DB 2; Length 1119;
Best Local Similarity 55.6%; Pred. No. 25;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 ORLKYAOELSENVLP 18
|| | | | | | | | | | |
Db 165 QROSYASEDYENEVSRPP 182

RESULT 14
T29757
protein UNC-89 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
C:Accession: T29757
R:Du, Z.; Le, T.T.; Wilson, R.
submitted to the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid C09D1.
A:Reference number: 220679
A:Accession: T29757
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-6642 <DUZ>
A:Cross-references: EMBL:AF003131; PIDN:AAB54132.1; GSPDB:GN00019; CESP:unc-89
A:Experimental source: strain Bristol N2; clone C09D1
C:Genetics:
A:Gene: CESP:unc-89
A:Map position: 1
A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1
/3; 5917/1; 6037/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

Query Match 39.5%; Score 51; DB 2; Length 6642;
Best Local Similarity 45.5%; Pred. NO. 1.9e+02;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 8.23529 Seconds
(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129
Sequence: 1 QRLKYAOEELSNEVLPPEPKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	341	1	P13672 hepatitis a
2	129	100.0	852	1	P06442 hepatitis a
3	129	100.0	2226	1	P06580 hepatitis a
4	129	100.0	2226	1	P26581 hepatitis a
5	129	100.0	2226	1	P26582 hepatitis a
6	129	100.0	2227	1	P06617 hepatitis a
7	129	100.0	2227	1	P06641 hepatitis a
8	123	95.3	2227	1	P13901 hepatitis a
9	119	92.2	2230	1	P14553 simian hepa
10	117	90.7	839	1	P031788 simian hepa
11	103	79.8	808	1	P032381 hepatitis a
12	53	41.1	636	1	Q15025 homo sapien
13	52	40.3	1193	1	Q10578 caenorhabdi
14	51	39.5	6632	1	U0899_CAEEL
15	50.5	39.1	443	1	HSIU_VIRCH
16	49.5	38.4	554	1	O8dcp4 vibrio chol
17	49.5	38.4	554	1	P51021 podospira a
18	49	38.0	1174	1	P30876 homo sapien
19	48.5	37.6	2278	1	P34756 saccharomyc
20	48	37.2	1176	1	P08266 diosiphila
21	48	37.2	1191	1	Q42877 lycopersico
22	48	37.2	1210	1	RPB2_LYCSC
23	47	36.4	578	1	P09765 schizosacch
24	47	36.4	592	1	P15891 saccharomyc
25	46.5	36.0	399	1	P34821 mus musculu
26	46	35.7	338	1	Q01664 homo sapien
27	46	35.7	480	1	P13744 cucurbita m
28	46	35.7	574	1	P03461 influenza b
29	46	35.7	576	1	P09766 influenza b
30	46	35.7	578	1	P09767 influenza b
31	46	35.7	583	1	P10757 influenza b
32	46	35.7	583	1	P03464 influenza b
33	46	35.7	583	1	P03463 influenza b

34	46	35.7	585	1	HEMA_INBE	P17504 influenza b
35	46	35.7	585	1	HEMA_INBYK	P22092 influenza b
36	46	35.7	1085	1	CARB_HELPJ	O9xkt2 helicobacte
37	45	34.9	315	1	RSEB_HAEIN	P44792 haemophilus
38	45	34.9	445	1	ML64_HUMAN	O14849 homo sapien
39	45	34.9	733	1	HEXA_BLADI	O17127 blaberus di
40	45	34.9	1188	1	RPB2_ARATH	P38420 arabidopsis
41	45	34.9	1312	1	L159_CAEEL	O44757 caenorhabdi
42	44.5	34.5	702	1	CRN_DROME	P17886 drosophila
43	44.5	34.5	3678	1	DMD_MOUSE	P11531 mus musculu
44	44.5	34.5	3680	1	DMD_CANFA	O97592 canis fam11
45	44.5	34.5	3685	1	DMD_HUMAN	P11532 homo sapien

ALIGNMENTS

RESULT 1
POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP3; Core protein
DE P2A) (Fragment).
OS Hepatitis A virus (strain CDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89263805; PubMed-2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV)."
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -!- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -!- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X14666; CA32794.1; -.
DR PIR: S04137; S04137.
KW POLYPROTEIN; Coat protein; Core protein.
FT NON_TER 1
FT CHAIN 1
FT CHAIN 2 340 COAT PROTEIN VP3 (1C).
FT CHAIN 341 >341 COAT PROTEIN VP1 (1D).
FT CHAIN 341 341 CORE PROTEIN P2A.
FT NON_TER 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 6,2e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 QRLKYAOEELSNEVLPPEPKMKGLF 25
DB 314 QRLKYAOEELSNEVLPPEPKMKGLF 338
RESULT 2
POLG_HPAVC STANDARD; PRT; 852 AA.
AC P06442; O83741; O83742;

DT 01-JAN-1988 (Rel. 06, Created)
DT 16-OCT-2001 (Rel. 06, Last sequence update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=815648; PubMed=2985793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RT Young A., Mitra S.W.,
RL "Molecular cloning and partial sequencing of hepatitis A viral cDNA."
J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PPM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M10033; AAA45470.1; -.
DR PIR: A03904; GNNYHA.
KW Polypeptide; Coat protein; Core protein.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 >852
FT NON_TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

Query Match 100.0%; Score 129; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSENYLPPPRKMKGLF 25
DB 810 QRLKYAOEELSENYLPPPRKMKGLF 834

ID POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Reinstone S.M.,
RT Cromeans T., Jansen R.W.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination."
J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
(RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PPM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59810; AAA45468.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR pfam: PF00680; RNA_dep_RNA_pol; 1.
DR pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICIVIRUSNS.
KW RNA-protein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSENYLPPPRKMKGLF 25
DB 810 QRLKYAOEELSENYLPPPRKMKGLF 834

ID POLG_HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Reinstone S.M.,
RT Cromeans T., Jansen R.W.,
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination."
J. Virol. 65:2056-2065(1991).

```
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59809: AAA45469.1; -.
CC DR MEROPS: C03.005; -.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVlr.
CC DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
CC DR Pfam: PF00910; RNA_helicase; 1.
CC DR Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC FT CHAIN 1 23
CC FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 795 900 COAT PROTEIN VP1 (PID).
CC FT CHAIN 901 1087 CORE PROTEIN P2A.
CC FT CHAIN 1088 1422 CORE PROTEIN P2B.
CC FT CHAIN 1423 1495 CORE PROTEIN P2C.
CC FT CHAIN 1496 1518 PROBABLE PROTEIN P3A.
CC FT CHAIN 1519 1737 PROBABLE PROTEIN P3B.
CC FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
CC FT CHAIN 2226 251107 MW; 403B4CA80B09BE75 CRC64;
CC SQ SEQUENCE
CC
CC Query Match 100.0%; Score 129; DB 1; Length 2226;
CC Best Local Similarity 100.0%; Pred. No. 4.9e-11;
CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1 QRLKYAOEELSNEVLPPPRKMKGLF 25
CC DB 810 QRLKYAOEELSNEVLPPPRKMKGLF 834
CC
CC RESULT 5
CC POLG_HPAV8 STANDARD; PRT; 2226 AA.
CC ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
CC AC P26582;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
CC P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
CC P3D (EC 2.7.7.48)].
CC OS Hepatitis A virus (strain 18f).
CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12096;
CC OX NCBI_TaxID=12096;
CC OX [1]
CC RN SEQUENCE FROM N.A.
CC RP MEDLINE=91162758; PubMed=1705995;
CC RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
CC Cremons T., Jansen R.W.;
CC "Antigenic and genetic variation in cytopathic hepatitis A virus
CC variants arising during persistent infection: evidence for genetic
CC recombination.";
CC RT J. Virol. 65:2056-2065(1991).
CC RL J. Virol. 65:2056-2065(1991).
CC CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
```

```
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59808: AAA45467.1; -.
CC DR PDB: 1OAT; 15-MAY-00.
CC DR MEROPS: C03.005; -.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVlr.
CC DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
CC DR Pfam: PF00910; RNA_helicase; 1.
CC DR Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
CC FT CHAIN 1 23
CC FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).
CC FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).
CC FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).
CC FT CHAIN 795 900 COAT PROTEIN VP1 (PID).
CC FT CHAIN 901 1087 CORE PROTEIN P2A.
CC FT CHAIN 1088 1422 CORE PROTEIN P2B.
CC FT CHAIN 1423 1495 CORE PROTEIN P2C.
CC FT CHAIN 1496 1518 PROBABLE PROTEIN P3A.
CC FT CHAIN 1519 1737 PROBABLE PROTEIN P3B.
CC FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.
CC FT CHAIN 2226 251292 MW; 24964A6396C8D6B CRC64;
CC SQ SEQUENCE
CC
CC Query Match 100.0%; Score 129; DB 1; Length 2226;
CC Best Local Similarity 100.0%; Pred. No. 4.9e-11;
CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC 1 QRLKYAOEELSNEVLPPPRKMKGLF 25
CC DB 810 QRLKYAOEELSNEVLPPPRKMKGLF 834
CC
CC RESULT 6
CC POLG_HPAVH STANDARD; PRT; 2227 AA.
CC ID POLG_HPAVH STANDARD; PRT; 2227 AA.
CC AC P08617; P06443; O81082;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-AUG-1988 (Rel. 08, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
CC P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
CC P3D (EC 2.7.7.48)].
CC OS Hepatitis A virus (strain HM-175).
CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12098;
CC OX NCBI_TaxID=12098;
CC OX [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=Wild type;
CC RX MEDLINE=87061253; PubMed=3023706;
CC RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
CC Baroudy B.M.;
CC "Complete nucleotide sequence of wild-type hepatitis A virus:
CC comparison with different strains of hepatitis A virus and other
CC picornaviruses.";
CC RT J. Virol. 61:50-59(1987).
CC RL [2]
CC RN SEQUENCE FROM N.A.
```

RC STRAIN-Attenuated;
 RA MEDLINE-87175701; PubMed-3031686;
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Felstone S.M.,
 RA Purcell R.H.;
 RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
 RT comparison with wild-type virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 RN [1]
 RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RX MEDLINE-85166289; PubMed-2984684;
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Matzel J.V. Jr.,
 RA Purcell R.H., Felstone S.M.;
 RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
 RT proteins and RNA polymerase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
 CC SHOWN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M14114; AAA45475.1; -;
 CC EMBL: M14707; AAA45465.1; -;
 CC EMBL: M14707; AAA45466.1; ALT_INTT.
 CC EMBL: M16632; AAA45471.1; -;
 CC PIR: A25981; GNNYMK.
 CC PIR: A94149; GNNYMK.
 CC DR PDB: 1HAV; 23-DEC-96.
 CC MEROPS: C03.005; -;
 CC InterPro: IPR004004; Callc1_pol_hel.
 CC InterPro: IPR000605; RNA_helicase.
 CC InterPro: IPR007095; RNA_pol_DS_PS.
 CC InterPro: IPR001205; RNA_pol_P3D.
 CC InterPro: IPR007094; RNA_pol_PSVlr.
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam: PF00910; RNA_helicase; 1.
 CC PRINTS: PR00918; CALICIVIRUSNS.
 CC -----
 CC KW Polypeptide; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 CC FT CHAIN 1 23
 CC FT CHAIN 24 245
 CC FT CHAIN 246 491
 CC FT CHAIN 492 836
 CC FT CHAIN 837 980
 CC FT CHAIN 981 1087
 CC FT CHAIN 1088 1422
 CC FT CHAIN 1423 1496
 CC FT CHAIN 1497 1519
 CC FT CHAIN 1520 1738
 CC FT CHAIN 1739 2227
 CC FT CHAIN 77 77
 CC FT CHAIN 764 764
 CC FT CHAIN 821 821
 CC FT CHAIN 1052 1052
 CC FT CHAIN 1052 1062
 CC FT CHAIN 1062 1118
 CC FT CHAIN 1118 1151
 CC FT CHAIN 1151 1163
 CC FT CHAIN 1163 1277
 CC FT CHAIN 1277 1277
 CC V -> I (IN ATTENUATED STRAIN).
 CC V -> S (IN ATTENUATED STRAIN).
 CC V -> K (IN ATTENUATED STRAIN).
 CC E -> K (IN ATTENUATED STRAIN).
 CC E -> S (IN ATTENUATED STRAIN).
 CC E -> I (IN ATTENUATED STRAIN).
 CC E -> V (IN ATTENUATED STRAIN).
 CC N -> S (IN ATTENUATED STRAIN).
 CC A -> V (IN ATTENUATED STRAIN).
 CC G -> A (IN ATTENUATED STRAIN).
 CC K -> M (IN ATTENUATED STRAIN).
 CC E -> K (IN ATTENUATED STRAIN).
 CC F -> S (IN ATTENUATED STRAIN).
 CC V -> I (IN ATTENUATED STRAIN).

FT VARIANT 1500 1500 H -> Y (IN ATTENUATED STRAIN).
 FT VARIANT 1805 1805 D -> N (IN ATTENUATED STRAIN).
 FT VARIANT 1930 1930 S -> T (IN ATTENUATED STRAIN).
 SQ SEQUENCE 2227 AA; 251506 MW; 01E23E74EB740A6 CRC64;
 Query Match 100.0%; Score 129; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4,9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 ORLKYAOEISNEVLPPEPRKMKGLF 25
 Db 810 ORLKYAOEISNEVLPPEPRKMKGLF 834
 RESULT 7
 POLG_HPAVL STANDARD; PRT; 2227 AA.
 ID POLG_HPAVL
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3A (RC 2.7.7.48)].
 OS Hepatitis A virus (strain 1A).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12099;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85190549; PubMed-2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
 RA Merryweather J., van Nest G., Dina D.;
 RT "Primary structure and gene organization of human hepatitis A virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: K02990; AAA45472.1; -;
 CC PIR: A03903; GNNYMK.
 CC MEROPS: C03.005; -;
 CC InterPro: IPR004004; Callc1_pol_hel.
 CC InterPro: IPR000605; RNA_helicase.
 CC InterPro: IPR007095; RNA_pol_DS_PS.
 CC InterPro: IPR001205; RNA_pol_P3D.
 CC InterPro: IPR007094; RNA_pol_PSVlr.
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam: PF00910; RNA_helicase; 1.
 CC PRINTS: PR00918; CALICIVIRUSNS.
 CC KW Polypeptide; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 CC FT CHAIN 1 23
 CC FT CHAIN 24 245
 CC FT CHAIN 246 491
 CC FT CHAIN 492 836
 CC FT CHAIN 837 980
 CC FT CHAIN 981 1076
 CC FT CHAIN 1077 1422
 CC FT CHAIN 1423 1484
 CC FT CHAIN 1485 1507
 CC FT CHAIN 1507 1507
 CC V -> I (IN ATTENUATED STRAIN).
 CC V -> S (IN ATTENUATED STRAIN).
 CC V -> K (IN ATTENUATED STRAIN).
 CC E -> K (IN ATTENUATED STRAIN).
 CC E -> S (IN ATTENUATED STRAIN).
 CC E -> I (IN ATTENUATED STRAIN).
 CC E -> V (IN ATTENUATED STRAIN).
 CC N -> S (IN ATTENUATED STRAIN).
 CC A -> V (IN ATTENUATED STRAIN).
 CC G -> A (IN ATTENUATED STRAIN).
 CC K -> M (IN ATTENUATED STRAIN).
 CC E -> K (IN ATTENUATED STRAIN).
 CC F -> S (IN ATTENUATED STRAIN).
 CC V -> I (IN ATTENUATED STRAIN).

FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
 FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 4.9e-11;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYA0EELSNEVLPPPRKMGLEF 25
 ||||||||||||||||||
 DB 810 ORLKYA0EELSNEVLPPPRKMGLEF 834

RESULT 8

POLG_HPAVM STANDARD; PRT: 2227 AA.
 ID POLG_HPAVM 081083; 081084; 081085; 081086; 081087; 081088; 081089;
 AC 081090; 081091; 081092; 081093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain AGM-27).
 OS Hepatitis A virus (strain AGM-27).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
 OX NCBI_TaxID-12102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88045071; PubMed-2823500;
 RA Paul A.V., Tada H., der Helm K., Wessel T., Klein R., Wimmer E.,
 RA Delhardt F.;
 RT "The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB).";
 RL Virus Res. 8:153-171(1987).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + {RNA}(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M20273; AAA5474.1; -
 DR MEROPS: C03.005; -
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_PSVIR.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease;
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1432
 FT CHAIN 1433 1496
 FT CHAIN 1497 1519
 PROBABLE PROTEIN P3A.
 PROBABLE PROTEIN P3B.

FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 95.3%; Score 123; DB 1; Length 2227;
 Best Local Similarity 96.0%; Pred. No. 3.9e-10;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ORLKYA0EELSNEVLPPPRKMGLEF 25
 ||||||||||||||||||
 DB 810 ORLKYA0EELSNEVLPPPRKMGLEF 834

RESULT 9

POLG_HPAVS STANDARD; PRT: 2230 AA.
 ID POLG_HPAVS 081083; 081084; 081085; 081086; 081087; 081088; 081089;
 AC 081090; 081091; 081092; 081093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Simian hepatitis A virus (strain AGM-27).
 OS Simian hepatitis A virus (strain AGM-27).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
 OX NCBI_TaxID-12102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91311420; PubMed-1649901;
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
 RA Purcell R.H.;
 RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";
 RL J. Gen. Virol. 72:1677-1683(1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + {RNA}(N).
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D00924; BAA00766.1; -
 DR EMBL: X15461; CAA33490.1; -
 DR PIR: A30470; GNNYSA.
 DR MEROPS: C03.005; -
 DR InterPro: IPR004004; Calic1_pol_hel.
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR007094; RNA_pol_PSVIR.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 DR PRINTS: PR00918; CALICVIRUSNS.
 DR Polypeptide; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.

```
FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 92.2%; Score 119; DB 1; Length 2230;
Best Local Similarity 88.0%; Pred. No. 1.6e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKKGLF 25
||||| ||||| ||||| ||||| |||||
814 ORLKYAOEELSNEVLPPPKKGLF 838

POLG_HPAVT STANDARD; PRT; 839 AA.
ID P31768;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 to VP4; Core protein
P2A] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=31707;
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
cytomegalus macaques (Macaca fascicularis).";
J. Gen. Virol. 72:1685-1689(1991).
RL -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59286; AAA5473.1; -
PIR: J01180; GNMX52.
KW Polypeptide; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E912DBC CRC64;

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKKGLF 25
||||| ||||| ||||| ||||| |||||
809 ORLKYAOEELSNEVLPPPKKGLF 833
```

```
RESULT 11
POLG_HPAVT STANDARD; PRT; 808 AA.
ID Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 to VP4; Core protein
P2A] (Fragment).
OS Hepatitis A virus (strain GA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=31706;
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Speldring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
J. Med. Virol. 36:118-124(1992).
RL -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M66695; AAA5477.1; -
KW Polypeptide; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN 1 2 COAT PROTEIN VP4 (PIA).
FT CHAIN 3 223 COAT PROTEIN VP2 (PIB).
FT CHAIN 224 470 COAT PROTEIN VP3 (PIC).
FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
FT CHAIN 771 >808 CORE PROTEIN P2A.
FT NON_TER 808 808
SQ SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;

Query Match 79.8%; Score 103; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPK 20
||||| ||||| ||||| ||||| |||||
789 ORLKYAOEELSNEVLPPPK 808

RESULT 12
NAFL_HUMAN STANDARD; PRT; 636 AA.
ID Q15025; O76008; Q96EL9; Q99833; Q9HJL3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nef-associated factor 1 (Naf1) (HIV-1 Nef interacting protein)
DE (Vifion-associated nuclear shuttling protein) (VAN) (hVAN) (TNPAIP3
interacting protein 1).
GN TNIP1 OR NAF1 OR KINA0113.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RX MEDLINE=99120485; PubMed=9923610;
SQ SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
```


RA Fukushimi M., Dixon J., Kimura T., Tsurutani N., Dixon M.J.,
RT "Identification and cloning of a novel cellular protein Nat1, Netf-
RT associated factor 1, that increases cell surface CD4 expression",
RL FEBS Lett. 442:83-88(1999).
RN (2)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Leukocyte;
EX MEDLINE=20541981; PubMed=11090181;
RA Gupta K., Ott D., Hope T.J., Siliciano R.F., Boeke J.D.;
RT "A human nuclear shuttling protein that interacts with human
RT immunodeficiency virus type 1 matrix is packaged into virions.",
RL J. Virol. 74:11811-11824(2000).
RN (3)
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Lung;
EX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derje J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Bluetel K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Sisle F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stempelson M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Tothylvaki S., Cantinici P., Prange C.,
RA Baha S.S., Loguclano N.A., peters G.J., Abramson R.D., Mulhaly S.J.,
RA Bosak S.A., McEwan P.J., McKenna R.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.M., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.",
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (4)
RP SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).
RC TISSUE=Craniofacial;
EX MEDLINE=96276047; PubMed=8681136;
RA Loebus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;
RT "Transcriptional map of the Treacher Collins candidate gene region.",
RL Genome Res. 6:26-34(1996).
RN (5)
RP SEQUENCE OF 341-636 FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow;
EX MEDLINE=95308325; PubMed=7788527;
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,
RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. III
RT The coding sequences of 40 new genes (K1AA0081-K1AA0120) deduced by
RT analysis of cDNA clones from human cell line KG-1.",
RL DNA Res. 2:37-43(1995).
RN (6)
RP SEQUENCE OF 94-412 FROM N.A.
RA Fukushimi M., Kimura T., Yamamoto N.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN (7)
RP FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and
RP inhibits TNF-induced NF-kappa-B-dependent gene expression by
RP interfering with an RIP- or TRAF2-mediated transactivation signal
RP (by similarity). Increases cell surface CD4(74) antigen
RP expression. Interacts with HIV-1 matrix protein and is packaged
RP into virions and overexpression can inhibit viral replication. May
RP regulate matrix nuclear localization, both nuclear import of PIC
RP (preintegration complex) and export of Gag polypeptide and viral
RP genomic RNA during virion production.
RN (8)
RP SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with
RP HIV-1 matrix protein.
RN (9)
RP SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus
RP and cytoplasm in a CRM1-dependent manner.
RN (10)
RP ALTERNATIVE PRODUCTS:
RP Event-Alternative splicing; Named isoforms=2;

```

CC      Name-1: Synonyms=Alpha:
CC      Isoid=Q15025-1; Sequence=Displayed:
CC      Name-2; Synonyms=Beta:
CC      Isoid=Q15025-2; Sequence=VSP_003913;
CC      Note=No experimental confirmation available;
CC      TISSUE SPECIFICITY: Ubiquitous. Strongly expressed in peripheral
CC      blood lymphocytes, spleen and skeletal muscle, and is weakly
CC      expressed in the brain.
CC      -1 CAUTION: Ref. 6 sequence differs from that shown due to frameshifts
CC      in positions 157 and 154.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      DR EMBL; AJ011895; CAAG0855.1; -
CC      DR EMBL; AJ011896; CAAG0856.1; -
CC      DR EMBL; AY012155; AAC42154.1; -
CC      DR EMBL; BC012133; AAH12133.1; -
CC      DR EMBL; BC014008; AAH14008.1; -
CC      DR EMBL; U39403; AAC99999.1; -
CC      DR EMBL; D30755; BAA0416.1; -
CC      DR EMBL; U83844; AAB41438.1; ALT_FRAME.
CC      DR Genew; HGNC:16903; TNIP1.
CC      DR GO; GO:0005515; F:protein binding activity; TAS.
CC      DR GO; GO:0009101; P:glycoprotein biosynthesis; IDA.
CC      DR GO; GO:0045071; P:negative regulation of viral genome replica. .; TAS.
CC      KM Colled coil; Nuclear protein; Alternative splicing.
CC      FT DOMAIN 20 73
CC      FT DOMAIN 196 258
CC      FT DOMAIN 294 535
CC      FT DOMAIN 94 412
CC      FT DOMAIN 524 530
CC      FT DOMAIN 539 636
CC      FT VARSPLIC 627 636
CC      FT CONFLICT 148 148
CC      FT CONFLICT 299 299
CC      SQ SEQUENCE 636 AA: 71864 MW: DB1B9EBAD50D871 CRC64;
CC      G -> D (IN REF. 3; AAH12133).
CC      /FTid=VSP_003913.
CC      G -> P (IN REF. 2).
CC      SPKDRDEQP -> PADRLRLPRN (In isoform 2).
CC      Query Match 41.1%; Score 53; DB 1; Length 636;
CC      Best Local Similarity 50.0%; Pred. No. 3.3;
CC      Matches 13; Conservative 2; Mismatches 9; Indels 2; Gaps 1;
CC      Oy 2 RLKTAQDEL--SNEVLPPPRMKGLF 25
CC      Db 58 RLROKADELVNDNELPPSPSLCSF 83
CC      II: III IIIIIII I
CC      RESULT 13
CC      RPB2_CADEL STANDARD: PRT; 1193 AA.
CC      ID RPB2_CADEL
CC      AC Q10578;
CC      DT 01-OCT-1996 (Rel. 34, Created)
CC      DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC      DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC      DE DNA-directed RNA polymerase II second largest subunit (EC 2.7.7.6)
CC      DE (RNA polymerase I subunit 2).
CC      GN C26E6.4.
CC      OS Caenorhabditis elegans.
CC      OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
CC      Rhabditidae; Pelodidae; Caenorhabditis.
CC      OX NCBI_Taxid=6239;
CC      RN (1)
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN-Bristol N2;
CC      RA Fulton L.;
CC      TL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC      [2]

```

RP SEQUENCE OF 200-1058 FROM N.A.
RX MEDLINE-9504134; PubMed-7953533;
RA Sldow A., Thomas W.K.;
RT "A molecular evolutionary framework for eukaryotic model organisms.";
RL Curr. Biol. 4:596-603(1994).
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA(N)).
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE SECOND LARGEST
CC COMPONENT OF RNA POLYMERASE II.
CC -1- SUBCELLULAR LOCATION: Nucleus.
CC -1- FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE
CC III FOR 5S AND TRNA GENES.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U13875; AAA21158.1; -
DR EMBL: U10333; AAA50224.1; -
DR PIR: E88445; E88445.
DR PIR: T43701; T43701.
DR WormPep: C2656.4; CE01162.
DR InterPro: IPR001572; RNA_POL_B.
DR Pfam: PF04563; RNA_POL_RPB2_1; 1.
DR Pfam: PF04561; RNA_POL_RPB2_2; 1.
DR Pfam: PF04565; RNA_POL_RPB2_3; 1.
DR Pfam: PF04566; RNA_POL_RPB2_4; 1.
DR Pfam: PF04567; RNA_POL_RPB2_5; 1.
DR Pfam: PF00562; RNA_POL_RPB2_6; 1.
DR Pfam: PF04560; RNA_POL_RPB2_7; 1.
DR PROSITE: PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein. C4-TYPE (POTENTIAL).
FT ZN.FING 1125 1146
FT SEQUENCE 1193 AA: 134904 MW: 8848574E9C7EBE CRC64;
SQ
Query Match 40.3%; Score 52; DB 1; Length 1193;
Best Local Similarity 62.5%; Pred. No. 9.3;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 1 ORLKYAGELSNVLP 16
II:||||:| | 1:11
DB 340 ORIKYAREILOKELLP 355
RESULT 14
UN89_CAEEL STANDARD; PRT: 6632 AA.
AC 001761; 017362;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).
GN UNC-89 OR C09D1.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
ON [1]
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RC STRAIN-Bristol N2;
RX MEDLINE-96180278; PubMed-8603916;

RA Benian G.M., Tinley T.L., Tang X., Borodovsky M.;
RT "The Caenorhabditis elegans gene unc-89, required for muscle M-line
RT assembly, encodes a giant modular protein composed of Ig and signal
RT transduction domains";
RL J. Cell Biol. 132:835-848(1996).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN-Bristol N2;
CC Du Z., Le T.T., Wilson R.;
CC Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
CC [3]
CC REVISIONS.
CC Waterston R.;
CC Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Structural component of the muscle M-line. Myofibril
CC lattice assembly begins with positional cues laid down in the
CC basement membrane and muscle cell membrane. UNC-89 responds to
CC these signals, localizes, and then participates in assembling an
CC M-line.
CC -1- TISSUE SPECIFICITY: Localizes to the middle of A-bands.
CC -1- SIMILARITY: Contains 1 DBL-homology (DH) domain.
CC -1- SIMILARITY: Contains 1 fibronectin type III domain.
CC -1- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 5 RCSD domains.
CC -1- SIMILARITY: Contains 1 SH3 domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U33058; AAB00542.1; -
DR EMBL: AF003131; AAB54132.2; -
DR PDB: 1FHC; 2O-DEC-00.
DR WormPep: C09D1.1; CE30426.
DR InterPro: IPR003961; FN_III.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001845; PH.
DR InterPro: IPR007850; RCSD.
DR InterPro: IPR000219; RhGEF.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00041; fn3; 1.
DR Pfam: PF00047; Ig; 47.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF05177; RCSD; 5.
DR Pfam: PF00621; RhGEF; 1.
DR Pfam: PF00018; SH3; 1.
DR SMART: SM00408; IGC2; 23.
DR SMART: SM00325; RhGEF; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50010; DH_2; 1.
DR PROSITE: PS50835; IG-LIKE; 49.
DR PROSITE: PS50003; PH DOMAIN; 1.
DR PROSITE: PS50002; SH3; 1.
KW Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
KW 3D-structure.
FT DOMAIN 63 127 SH3.
FT DOMAIN 152 330 DH.
FT DOMAIN 342 498 PH.
FT DOMAIN 547 633 IG-LIKE C2-TYPE 1.
FT DOMAIN 648 736 IG-LIKE C2-TYPE 2.
FT DOMAIN 748 838 IG-LIKE C2-TYPE 3.
FT DOMAIN 946 1033 IG-LIKE C2-TYPE 4.
FT DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.
FT DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.
FT DOMAIN 1272 1315 THR-RICH.
FT DOMAIN 1375 1475 RCSD 1.

```

FT DOMAIN 1479 1585 RCSD 2.
FT DOMAIN 1597 1695 RCSD 3.
FT DOMAIN 1700 1799 RCSD 4.
FT DOMAIN 1800 1860 RCSD 5.
FT DOMAIN 1982 2067 IG-LIKE C2-TYPE 7.
FT DOMAIN 2071 2163 IG-LIKE C2-TYPE 8.
FT DOMAIN 2171 2263 IG-LIKE C2-TYPE 9.
FT DOMAIN 2269 2359 IG-LIKE C2-TYPE 10.
FT DOMAIN 2367 2455 IG-LIKE C2-TYPE 11.
FT DOMAIN 2463 2564 IG-LIKE C2-TYPE 12.
FT DOMAIN 2563 2651 IG-LIKE C2-TYPE 13.
FT DOMAIN 2657 2746 IG-LIKE C2-TYPE 14.
FT DOMAIN 2754 2858 IG-LIKE C2-TYPE 15.
FT DOMAIN 2867 2964 IG-LIKE C2-TYPE 16.
FT DOMAIN 2994 3081 IG-LIKE C2-TYPE 17.
FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.
FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.
FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.
FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.
FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.
FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.
FT DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.
FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
FT DOMAIN 4106 4201 IG-LIKE C2-TYPE 27.
FT DOMAIN 4201 4297 IG-LIKE C2-TYPE 28.
FT DOMAIN 4297 4387 IG-LIKE C2-TYPE 29.
FT DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.
FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.
FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.
FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.
FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.
FT DOMAIN 4771 4861 IG-LIKE C2-TYPE 35.
FT DOMAIN 4861 4965 IG-LIKE C2-TYPE 36.
FT DOMAIN 4965 5067 IG-LIKE C2-TYPE 37.
FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.
FT DOMAIN 5260 5366 IG-LIKE C2-TYPE 39.
FT DOMAIN 5366 5472 IG-LIKE C2-TYPE 40.
FT DOMAIN 5472 5578 IG-LIKE C2-TYPE 41.
FT DOMAIN 5578 5685 IG-LIKE C2-TYPE 42.
FT DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.
FT DOMAIN 5790 5885 IG-LIKE C2-TYPE 44.
FT DOMAIN 5885 5975 IG-LIKE C2-TYPE 45.
FT DOMAIN 5975 6014 IG-LIKE C2-TYPE 46.
FT DOMAIN 6014 6130 IG-LIKE C2-TYPE 47.
FT DOMAIN 6130 6239 IG-LIKE C2-TYPE 48.
FT DOMAIN 6239 6368 FIBRONECTIN TYPE-III.
FT DOMAIN 6368 6413 IG-LIKE C2-TYPE 49.
FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 49.
FT DOMAIN 6502 6596 POTENTIAL.
FT DOMAIN 6596 6621 POTENTIAL.
FT DOMAIN 6621 6715 POTENTIAL.
FT DOMAIN 6715 6805 POTENTIAL.
FT DOMAIN 6805 6890 POTENTIAL.
FT DOMAIN 6890 6926 POTENTIAL.
FT DOMAIN 6926 7032 POTENTIAL.
FT DOMAIN 7032 7157 POTENTIAL.
FT DOMAIN 7157 7298 POTENTIAL.
FT DOMAIN 7298 7369 POTENTIAL.
FT DOMAIN 7369 7464 POTENTIAL.
FT DOMAIN 7464 7591 POTENTIAL.
FT DOMAIN 7591 7664 POTENTIAL.
FT DOMAIN 7664 7722 POTENTIAL.
FT DOMAIN 7722 7836 POTENTIAL.
FT DOMAIN 7836 7946 POTENTIAL.
FT DOMAIN 7946 8036 POTENTIAL.
FT DOMAIN 8036 8171 POTENTIAL.
FT DOMAIN 8171 8261 POTENTIAL.
FT DOMAIN 8261 8388 A -> P (IN REF. 1).
FT DOMAIN 8388 8486 AKA -> PKP (IN REF. 1).
FT DOMAIN 8486 8598 A -> P (IN REF. 1).
FT DOMAIN 8598 8671 E -> G (IN REF. 1).
FT DOMAIN 8671 8728 M -> I (IN REF. 1).
FT DOMAIN 8728 8888 A -> G (IN REF. 1).
FT DOMAIN 8888 8929 DAGEY -> RRRRI (IN REF. 1).
FT DOMAIN 8929 9034 A -> V (IN REF. 1).
FT DOMAIN 9034 9145 A -> P (IN REF. 1).
FT DOMAIN 9145 9261 T -> S (IN REF. 1).
FT DOMAIN 9261 9385 G -> A (IN REF. 1).

```

```

FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 F -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SO SEQUENCE 6632 AA: 731665 MW: 262D3ED62560E89 CRC64;

Query Match
Best Local Similarity 39.5%; Score 51; DB 1; Length 6632;
Matches 10; Conservative 45.5%; Pred. No. 86; Mismatches 7; Indels 0; Gaps 0;

QY 1 ONLKAQDELSENVLPKPKM 22
DB 1353 RRVSPAEELPREVIDSRKK 1374

RESULT 15
HSLU_VIBCH STANDARD; PRT; 443 AA.
ID HSLU_VIBCH
AC 09KNQ7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent hsl protease ATP-binding subunit hslU.
GN HSLU OR VC2674.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Yamachyan J., Bass S., Qin H., Diegel I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Mierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- FUNCTION: CHAPERONE SUBUNIT OF A PROTEASOME-LIKE DEGRADATION
CC COMPLEX (BY SIMILARITY).
CC -1- SUBUNIT: Interacts with hslU (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY. HSLU SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC
CC EMBL, AE004333; AAF95815.1;
CC PIR: E82046; E82046.
CC HSSP: P32168; IDO2.
CC
CC TIGR: VC2674;
CC
CC HAMAP: MF_00249;
CC
CC InterPro: IPR003593; AAA_Arpase.
CC InterPro: IPR003595; AAA_Arpase_cent.
CC InterPro: IPR004491; Hsp_HslU.
CC Pfam: PF00004; AAA; 1.
CC SMART: SM00382; AAA; 1.
CC TIGRfams: TIGR00390; hslU; 1.
CC Chaperone; ATP-binding; Complete proteome.
KW NP_BIND 57 64 ATP (POTENTIAL).
SO SEQUENCE 443 AA: 49900 MW: DAE13E2F8B6A38F CRC64;

Query Match 39.1%; Score 50.5; DB 1; Length 443;

```

Best Local Similarity 39.3%; Pred. No. 5.3;
Matches 11; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

OY 1 ORLKVAOEELSN-----VLPPPRKMG 23
:::| | | : | | | | |
Db 117 EKVFRRAEELAEERVLDALPPPRDAG 144

Search completed: October 1, 2003, 09:57:55
Job time : 9.23529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 41.4706 Seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129

Sequence: 1 QRLRYAOEELSNEVLPPIPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Actual number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp-archaea:*
- 2: sp-bacteria:*
- 3: sp-fungi:*
- 4: sp-human:*
- 5: sp-invertebrate:*
- 6: sp-mammal:*
- 7: sp-mhc:*
- 8: sp-organelle:*
- 9: sp-phage:*
- 10: sp-plant:*
- 11: sp-rodent:*
- 12: sp-virus:*
- 13: sp-vertebrate:*
- 14: sp-unclassified:*
- 15: sp-virus:*
- 16: sp-bacteriophage:*
- 17: sp-archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	94	12 Q9ENT9	Q9ent9 hepatitis a
2	129	100.0	94	12 Q9ENT3	Q9ent3 hepatitis a
3	129	100.0	94	12 Q9ENV9	Q9env9 hepatitis a
4	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
5	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
6	129	100.0	94	12 Q9ENT1	Q9ent1 hepatitis a
7	129	100.0	94	12 Q9ENT8	Q9ent8 hepatitis a
8	129	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
9	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
10	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
11	129	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
12	129	100.0	94	12 Q9ENV9	Q9env9 hepatitis a
13	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
14	129	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
15	129	100.0	94	12 Q9ENV6	Q9env6 hepatitis a
16	129	100.0	94	12 Q9ENW0	Q9enw0 hepatitis a

17	129	100.0	94	12 Q9ENT6	Q9ent6 hepatitis a
18	129	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
19	129	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
20	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
21	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
22	129	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
23	129	100.0	115	12 Q9DWR4	Q9dwr4 hepatitis a
24	129	100.0	116	12 Q92941	Q92941 hepatitis a
25	129	100.0	116	12 Q9W7X7	Q9w7x7 hepatitis a
26	129	100.0	116	12 Q9W7S7	Q9w7s7 hepatitis a
27	129	100.0	116	12 Q71977	Q71977 hepatitis a
28	129	100.0	116	12 Q8B8K6	Q8b8k6 hepatitis a
29	129	100.0	116	12 Q8B8K5	Q8b8k5 hepatitis a
30	129	100.0	116	12 Q8B8K4	Q8b8k4 hepatitis a
31	129	100.0	116	12 Q8B8K3	Q8b8k3 hepatitis a
32	129	100.0	126	12 Q9BMT9	Q9bmt9 hepatitis a
33	129	100.0	132	12 Q8V4L9	Q8v4l9 hepatitis a
34	129	100.0	132	12 Q8V4L6	Q8v4l6 hepatitis a
35	129	100.0	132	12 Q8V4J9	Q8v4j9 hepatitis a
36	129	100.0	132	12 Q8V4J5	Q8v4j5 hepatitis a
37	129	100.0	132	12 Q8V4M5	Q8v4m5 hepatitis a
38	129	100.0	132	12 Q8V4M7	Q8v4m7 hepatitis a
39	129	100.0	132	12 Q8V4N2	Q8v4n2 hepatitis a
40	129	100.0	132	12 Q8V4M4	Q8v4m4 hepatitis a
41	129	100.0	132	12 Q8V4J7	Q8v4j7 hepatitis a
42	129	100.0	132	12 Q8V4L5	Q8v4l5 hepatitis a
43	129	100.0	132	12 Q8V4K3	Q8v4k3 hepatitis a
44	129	100.0	132	12 Q8V4I0	Q8v4i0 hepatitis a
45	129	100.0	132	12 Q8V4M2	Q8v4m2 hepatitis a

ALIGNMENTS

RESULT 1

ID Q9ENT9 PRELIMINARY: PRT: 94 AA.

AC Q9ENT9;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Nagasaki 32;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y., Yasuoka A., Oka S.;

RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV) infection: Prolonged HAV viremia and mild liver injury.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB036298; BAB1836.1; -

FT NON_TER 1

FT NON_TER 94

SO SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 4.4e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QRLRYAOEELSNEVLPPIPRKMKGLF 25

DB 61 QRLRYAOEELSNEVLPPIPRKMKGLF 85

RESULT 2

ID Q9ENV3 PRELIMINARY: PRT: 94 AA.

AC Q9ENV3;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

```
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
OS Polypotein (Fragment).
OC Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 27;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB038294; BAB11832.1; -.
FT NON_TER 1 1
FT SEQUENCE 94 AA; 10892 MW; 9AF8BE91BB8C53CC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 3
Q9ENV9 PRELIMINARY; PRT; 94 AA.
AC Q9ENV9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 03;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB038278; BAB11816.1; -.
FT NON_TER 1 1
FT SEQUENCE 94 AA; 10862 MW; 9AF9EFD4A8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 4
Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

```
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB038282; BAB11820.1; -.
FT NON_TER 1 1
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 5
Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 28;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB038295; BAB11833.1; -.
FT NON_TER 1 1
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 6
Q9ENV1 PRELIMINARY; PRT; 94 AA.
AC Q9ENV1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Kantou 49;
```

RA	Ide S., Tsuchikawa N., Nakajima A., Daitoku M., Yano M., Kikuchi Y.
RA	Yasuoka A., Oka S.:
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT	infection: Prolonged HAV viremia and mild liver injury.",
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL	EMBL: AB038306; BAB1844.1; -.
DR	
FT	NON_TER 1 1
FT	NON_TER 94 94
CO	SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;

Query Match	100.0%;	Score 129;	DB 12;	Length 94;
Best Local Similarity	100.0%;	Pred. No. 4.4e-12;		
Matches 25; Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

Oy 1 QRLKYAQEELSNVLPPrKMKGLF 25
 |||||
Db 61 QRLKYAQEELSNVLPPrKMKGLF 85

	SOLT	7
END1		
ID	OGENU1	PRELIMINARY:
AC	OGENU1;	PRT; 94 AA.
DT	01-MAR-2001 (TREMBLEI, 16, Created)	
DY	01-MAR-2001 (TREMBLEI, 16, Last sequence update)	
DT	01-MAR-2001 (TREMBLEI, 16, Last annotation update)	
DE	Polyprotein A (Fragment).	
OS	Hepatitis A virus.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;	
CC	Hepatovirus.	
OX	NCBI_TaxID=12092;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Nagaaki.29;	
RA	Iida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,	
RA	Yasuoka A., Oka S.;	
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)	
RT	infection: Prolonged HAV viremia and mild liver injury.";	
RL	Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.	
DR	EMBL; AB038296; BAB11834.1; .	
FT	NON_TER	
FT	NON_TER	
QO	SEQUENCE 94 AA, 10876 MW, 9AF8BE91BB8CABAC CRC64;	

Query Match	100.0%	Score 129;	DB 12;	Length 94;
Best Local Similarity	100.0%	Pred. No. 4.4e-12;		
Matches 25; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Db

1 QRLKYAQEELSNVLPPrKMKGLF 25
|||||
61 QRLKYAQEELSNVLPPrKMKGLF 85

RESULT 8
Q9EN08
ID Q9EN08 PRELIMINARY; PRT; 94 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis A virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
 CC Hepatovirus.
 OX NCBI_Taxid=12092;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN-Nagaasaki 22;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.;
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RT infection: Prolonged HAV viremia and mild liver injury.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR	EMBL, AB038289; BAB11827.1; -.
FT	NON_TER 1
FT	NON_TER 94
SQ	SEQUENCE 94 AA: 10876 MW: 9A9BBE91BB8C4BAC CRC64

Query Match	100.0%	Score 129, DB 12	Length 94,
Best Local Similarity	100.0%	Pred. No. 4, 4e-12	
Matches 25, Conservative	0	Mismatches 0;	Indels 0; Gaps 0

```

Qy      1 QRLKYADEELSNVLPPrPKMKGLF 25
          |||||
Db      61 QRLKYADEELSNVLPPrPKMKGLF 85

```

RESULT 9			
ID	09ENV2	PRELIMINARY:	PRF: 94 AA.
AC	09ENV2		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)		
DE	Polypotein (Fragment).		
OS	Hepatitis A virus.		
CC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
CC	Hepatovirus.		
OX	NCBI_TaxID=12092;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Nagasaki.15;		
RA	Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,		
RT	Yasuoka A., Oka S.;		
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)		
RL	infection: Prolonged HAV viremia and mild liver injury.";		
DR	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
EMBL:	AB038285; BAB11823.1; -		
FT	NON_TER 1 1		
FT	NON_TER 94 94		
SO	SEQUENCE 94 AA; 10876 MW; 9AF8BEF1BB8C4BAC CRC64;		

Query Match	100.0%	Score 139,	DB 12,	Length 94,
Best Local Similarity	100.0%	Pred. NC, 4.4e-12,		
Matches 25, Conservative	0,	Mismatches 0,	Indels 0,	Gaps 0

```
QY      1 QRLKAAQEELSNVLPPrPKMKGLF 25
        |||||
Db      61 QRLKAAQEELSNEVLPPPrKMKGLE 85
```

RESULT 10	
Q9ENUS	
ID	Q9ENUS
Q9ENUS	
PRELIMINARY;	PRT;
	94 AA

DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
DE Hepatitis A virus.
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae.
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX NCBI
RN SEQUENCE FROM N.A.
RP STRAIN=Nagasaki 25;
RC Iida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB038292; BAB11830.1; -.
FT NON_TER 1
FT NON_TER 94
FT SEQUENCE 94 AA: 10876 MW: 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
Db 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 11

Q9ENV1 PRELIMINARY; PRT; 94 AA.
AC O9ENV1; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;

SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 18;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038286; BAB11824.1; -.
FT NON_TER 1
FT 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
Db 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 12

Q9ENV9 PRELIMINARY; PRT; 94 AA.
AC O9ENV9; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;

SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 21;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038288; BAB11826.1; -.
FT NON_TER 1
FT 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPrKKKGLF 25

Db 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 13

Q9ENV7 PRELIMINARY; PRT; 94 AA.
AC O9ENV7; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;

SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 06;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038280; BAB11818.1; -.
FT NON_TER 1
FT 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
Db 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 14

Q9ENV8 PRELIMINARY; PRT; 94 AA.
AC O9ENV8; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
NCBI_TaxID=12092;

SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 04;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038279; BAB11817.1; -.
FT NON_TER 1
FT 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPrKKKGLF 25
Db 61 ORLKYAOEELSNEVLPPrKKKGLF 85

RESULT 15

ID	Q9ENV6	PRELIMINARY:	PRT:	94 AA.
AC	Q9ENV6			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	01-MAR-2001 (TREMBLrel. 16, Last annotation update)			
DE	Polyprotein (Fragment).			
OS	Hepatitis A virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.			
OX	NCBI_TaxID=12092;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Nagasaki 07;			
RA	Ida S.;			
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV) infection: Prolonged HAV viremia and mild liver injury."			
RL	Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AB038281; BAB11819.1; -.			
	NON_TER 1			
	NON_TER 94			
SEQUENCE	94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;			
Query Match	100.0%;	Score 129;	DB 12;	Length 94;
Best Local Similarity	100.0%;	Pred. No. 4.4e-12;		
Matches	25; Conservative	0; Mismatches	0; Indels	0; Gaps
QY	1 ORLKYAOEELSENYLPPRRKKMGFLF 25			
DB	61 ORLKYAOEELSENYLPPRRKKMGFLF 85			

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 15.8824 seconds
(without alignments)
66.600 Million cell updates/sec

Title: US-09-171-432a-47
Perfect score: 129
Sequence: 1 QRLKYAOEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	1091	6	5516630-2 Patent No. 5516630
2	129	100.0	2227	3	US-08-475-886-2 Sequence 2, Appl
3	129	100.0	2227	3	US-08-475-886-6 Sequence 6, Appl
4	129	100.0	2227	3	US-08-397-232-2 Sequence 2, Appl
5	129	100.0	2227	3	US-08-397-232-4 Sequence 4, Appl
6	129	100.0	2227	3	US-09-171-387-2 Sequence 2, Appl
7	129	100.0	2227	4	US-09-653-499-6 Sequence 2, Appl
8	129	100.0	2227	4	US-08-475-886-4 Sequence 6, Appl
9	124	96.1	2227	4	US-09-653-499-4 Sequence 4, Appl
10	124	96.1	2227	4	US-08-475-886-4 Sequence 4, Appl
11	117	90.7	839	4	US-08-087-016-2 Sequence 2, Appl
12	49	38.0	686	4	US-09-328-352-4303 Sequence 4303, Ap
13	47	36.4	180	4	US-09-252-991A-18339 Sequence 18339, A
14	46.5	36.0	23	1	PCT-US91-03368-4 Sequence 4, Appl
15	46.5	36.0	23	1	US-08-278-729A-8 Sequence 8, Appl
16	46.5	36.0	139	1	US-08-155-343A-8 Sequence 8, Appl
17	46.5	36.0	139	1	US-08-406-672-8 Sequence 8, Appl
18	46.5	36.0	139	1	US-08-643-563A-8 Sequence 8, Appl
19	46.5	36.0	139	1	US-08-643-563A-8 Sequence 8, Appl
20	46.5	36.0	139	1	US-08-643-763A-8 Sequence 8, Appl
21	46.5	36.0	139	1	US-08-462-623-8 Sequence 8, Appl
22	46.5	36.0	139	1	US-08-451-953A-8 Sequence 8, Appl
23	46.5	36.0	139	2	US-08-445-468A-8 Sequence 8, Appl
24	46.5	36.0	139	2	US-08-461-397A-8 Sequence 8, Appl
25	46.5	36.0	139	2	US-08-912-088-8 Sequence 8, Appl
26	46.5	36.0	139	3	US-08-278-730A-8 Sequence 8, Appl
27	46.5	36.0	139	3	US-08-445-467-8 Sequence 8, Appl

28	46.5	36.0	139	3	US-08-480-515A-8 Sequence 8, Appl
29	46.5	36.0	139	3	US-08-414-033A-8 Sequence 8, Appl
30	46.5	36.0	139	3	US-08-271-556A-6 Sequence 6, Appl
31	46.5	36.0	139	3	US-08-440-894A-8 Sequence 8, Appl
32	46.5	36.0	139	4	US-09-170-936-8 Sequence 8, Appl
33	46.5	36.0	139	4	US-08-461-113-8 Sequence 8, Appl
34	46.5	36.0	139	4	US-08-456-033-8 Sequence 8, Appl
35	46.5	36.0	139	4	US-08-643-321-7 Sequence 7, Appl
36	46.5	36.0	139	4	US-08-464-206-8 Sequence 8, Appl
37	46.5	36.0	139	4	US-08-404-113A-8 Sequence 8, Appl
38	46.5	36.0	139	5	PCT-US92-01968-8 Sequence 8, Appl
39	46.5	36.0	139	5	PCT-US93-07150-8 Sequence 8, Appl
40	46.5	36.0	139	5	PCT-US93-07231-8 Sequence 8, Appl
41	46.5	36.0	139	5	PCT-US93-08742-8 Sequence 8, Appl
42	46.5	36.0	139	5	PCT-US93-08808-8 Sequence 8, Appl
43	46.5	36.0	139	5	PCT-US93-08885-8 Sequence 8, Appl
44	46.5	36.0	397	5	US-07-841-646-27 Sequence 27, Appl
45	46.5	36.0	397	5	PCT-US91-07635-4 Sequence 4, Appl

ALIGNMENTS

```

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; PEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2:
; LENGTH: 1091
5516630-2

Query Match      100.0%; Score 129; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 QRLKYAOEELSNEVLPPPRKMKGLF 25
DB      1047 QRLKYAOEELSNEVLPPPRKMKGLF 1071

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE D
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262052
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
    
```

SEQ ID NO 2
LENGTH: 2227
TYPE: PRF
ORGANISM: MILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1993-09-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

ULT 4
08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
EARLIER FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRF
ORGANISM: MILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
EARLIER FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
GENERAL INFORMATION:
APPLICANT: RAYCHAUDHURI, GOPA;
EMERSON, SUZANNE, U.;
PURCELL, ROBERT, H.
TITLE OF INVENTION: HAVING A CHIMERIC 2C PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997

```
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-42290S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-171-387-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMG...LF 25
DB      810 ORLKTAQDELSEVLP...PPRKMG...LF 834

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMG...LF 25
DB      810 ORLKTAQDELSEVLP...PPRKMG...LF 834

RESULT 8
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
```

```
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMG...LF 25
DB      810 ORLKTAQDELSEVLP...PPRKMG...LF 834

RESULT 9
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      96.1%; Score 124; DB 3; Length 2227;
Best Local Similarity 96.0%; Pred. No. 8.9e-11;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMG...LF 25
DB      810 ORLKTAQDELSEVLP...PPRKMG...LF 834

RESULT 10
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31.
```


TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/800,364B
FILING DATE: 26-NOV-1991
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,245
REFERENCE/DOCKET NUMBER: GI 5182A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-876-1170
TELEFAX: 617-876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Bone
US-07-800-364B-4

Query Match 36.0%; Score 46.5; DB 1; Length 23;
Best Local Similarity 60.0%; Pred. No. 0.84;
Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 11 SNEVLPPPKMKGLF 25
:||| ||| | :|||
Db 1 TNE-LPPNKLPGIF 14

SUITE 15
US91-03388-4
Sequence 4, Application PC/TUS9103388
GENERAL INFORMATION:
APPLICANT: Hewick, Rodney M.
TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/03388
FILING DATE: 19910515
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.

REGISTRATION NUMBER: 32,245
REFERENCE/DOCKET NUMBER: GI5182X-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-876-1170
TELEFAX: 617-876-5851
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Bone
PCT-US91-03388-4

Query Match 36.0%; Score 46.5; DB 5; Length 23;
Best Local Similarity 60.0%; Pred. No. 0.84;
Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 11 SNEVLPPPKMKGLF 25
:||| ||| | :|||
Db 1 TNE-LPPNKLPGIF 14

Search completed: October 1, 2003, 10:06:34
Job time : 16.8824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 30.1471 Seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-47
Perfect score: 129
Sequence: 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Final number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCV_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/CTUS_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/CTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	129	100.0	980	US-10-272-459-41	Sequence 41, Appl
3	129	100.0	2227	US-09-929-955-12	Sequence 12, Appl
4	129	100.0	2227	US-10-104-966-12	Sequence 12, Appl
5	129	100.0	2227	US-10-135-988-2	Sequence 2, Appl
6	129	100.0	2227	US-10-135-988-6	Sequence 6, Appl
7	129	100.0	836	US-10-272-459-40	Sequence 40, Appl
8	124	96.1	2227	US-10-135-988-4	Sequence 4, Appl
9	46.5	36.0	139	US-08-260-675-8	Sequence 8, Appl
10	46.5	36.0	139	US-10-050-050-8	Sequence 0, Appl
11	46.5	36.0	399	US-08-957-425-27	Sequence 27, Appl
12	46.5	36.0	399	US-08-260-675-23	Sequence 23, Appl
13	46.5	36.0	399	US-10-350-747-2	Sequence 2, Appl
14	46.5	36.0	399	US-10-122-026-8	Sequence 8, Appl
15	46.5	36.0	399	US-10-050-050-23	Sequence 23, Appl

16	46	35.7	263	9	US-09-864-761-37656	Sequence 37656, A
17	45.5	35.3	1596	11	US-09-909-5678-47	Sequence 47, Appl
18	45	34.9	223	10	US-09-895-913A-84	Sequence 84, Appl
19	45	34.9	445	15	US-10-177-293-296	Sequence 296, App
20	43.5	33.7	222	15	US-10-156-761-11163	Sequence 11163, A
21	43.5	33.7	1114	12	US-09-840-743-14	Sequence 14, Appl
22	43	33.3	311	15	US-10-156-761-10491	Sequence 10491, A
23	43	33.3	442	15	US-10-156-761-13833	Sequence 13833, A
24	43	33.3	592	12	US-10-100-294A-36	Sequence 36, Appl
25	43	33.3	869	9	US-09-815-242-5230	Sequence 5230, Ap
26	43	33.3	882	9	US-09-815-242-12526	Sequence 12526, A
27	43	33.3	1095	15	US-10-128-714-3039	Sequence 3039, Ap
28	43	33.3	1209	15	US-10-128-714-3144	Sequence 3144, Ap
29	43	33.3	1255	15	US-10-128-714-8144	Sequence 8144, Ap
30	43	33.3	1277	15	US-10-128-714-8039	Sequence 8039, Ap
31	43	33.3	1370	11	US-09-842-758-111	Sequence 111, App
32	43	33.3	1570	11	US-09-842-758-35	Sequence 35, Appl
33	43	33.3	1648	10	US-09-515-806-4	Sequence 4, Appl
34	43	33.3	1648	11	US-09-842-758-37	Sequence 37, Appl
35	43	33.3	1648	11	US-09-842-758-39	Sequence 39, Appl
36	43	33.3	2696	12	US-10-309-933-4	Sequence 4, Appl
37	42	32.6	88	9	US-09-864-761-35516	Sequence 35516, A
38	42	32.6	88	9	US-09-864-761-43148	Sequence 43148, A
39	42	32.6	93	10	US-09-925-300-1101	Sequence 1101, Ap
40	42	32.6	146	10	US-09-923-246-56	Sequence 56, Appl
41	42	32.6	146	10	US-09-825-561A-47	Sequence 47, Appl
42	42	32.6	146	15	US-10-295-723-56	Sequence 56, Appl
43	42	32.6	192	14	US-10-001-857-119	Sequence 119, App
44	42	32.6	223	15	US-10-205-823-101	Sequence 101, Appl
45	42	32.6	311	8	US-08-818-581B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHTANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
US-10-272-459-45

Query Match 100.0% Score 129, DB 15, Length 352:
Best Local Similarity 100.0% Pred. No. 1.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25
DB 182 ORLKYAOEELSNEVLPPEPRKMKGLF 206
RESULT 2
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHTANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```

; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI/955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272.459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match          100.0%; Score 129; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPEPRKMGLE 25
|||||
810 ORLKYAOEELSNEVLPPEPRKMGLE 834

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match          100.0%; Score 129; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPEPRKMGLE 25
|||||
810 ORLKYAOEELSNEVLPPEPRKMGLE 834

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPEPRKMGLE 25
|||||
810 ORLKYAOEELSNEVLPPEPRKMGLE 834

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPEPRKMGLE 25
|||||
810 ORLKYAOEELSNEVLPPEPRKMGLE 834

RESULT 6
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 129; DB 14;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKTAQEELSNEVLPPPRKMKGLF 25
DB 810 ORLKTAQEELSNEVLPPPRKMKGLF 834

RESULT 7

US-10-272-459-40
Sequence 40, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANTES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: PP17955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 40
LENGTH: 836
TYPE: PRF
ORGANISM: Artificial Sequence
FEATURE: OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
OTHER INFORMATION: of 94 kDa
US-10-272-459-40

Query Match 96.9%; Score 125; DB 15;
Best Local Similarity 96.0%; Pred. No. 1e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKTAQEELSNEVLPPPRKMKGLF 25
DB 810 ORLKTAQEELSNEVLPPPRKMKGLF 834

RESULT 8

US-10-135-988-4
Sequence 4, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262053
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 96.1%; Score 124; DB 14;
Best Local Similarity 96.0%; Pred. No. 4.2e-09;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKTAQEELSNEVLPPPRKMKGLF 25
DB 810 ORLKTAQEELSNEVLPPPRKMKGLF 834

RESULT 9

US-08-260-675-8
Sequence 8, Application US/08260675
Publication No. US20030104993A1
GENERAL INFORMATION:
APPLICANT: RUBGER, DAVID C
APPLICANT: KUBERSAMPATH, THANGAVEL
APPLICANT: OPPERMAN, HERMANN
APPLICANT: OZKAYNAK, ENGIN
APPLICANT: PANG, ROY HL
APPLICANT: COHEN, CHARLES M
TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
TITLE OF INVENTION: REPAIR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESSES:
ADDRESSEE: TESTA HURWITZ & THIBEAULT
STREET: 55 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/260,675
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/126,100
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/922,813
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/667,274
FILING DATE: 11-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/752,764
FILING DATE: 30-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: MURIDAE
TISSUE TYPE: EMBRYO
FEATURE:
NAME/KEY: Protein
LOCATION: 1..139
OTHER INFORMATION: /label= MOP2-MATURE

US-08-260-675-8

Query Match 36.0%; Score 46.5; DB 8; Length 139;
Best Local Similarity 47.8%; Pred. No. 35;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 3 LKYOELSNVLPPrKMGlf 25
DB 5 LKRQPKTKNE-LPHPNKLPgIf 26

RESULT 10

US-10-050-050-8
Sequence 0, Application US/10050050
Publication No. US20030125230A1

GENERAL INFORMATION:

APPLICANT: COHEN, CHARLES M.

CHARRETTE, MARC F.

KUBERASAMPATH, THANGAVEL

ROEGER, DAVID C.

OPPERMANN, HERMANN

PANG, ROY H.L.

OZKAYNAK, ENGIN

SMART, JOHN E.

TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING

PROLIFERATION OF EPITHELIAL CELLS.

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES

STREET: 45 SOUTH STREET

CITY: HOPKINTON

STATE: MA

COUNTRY: USA

ZIP: 01748

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/050,050

FILING DATE: 15-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/461,113

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/445,882

FILING DATE: 22-MAY-1995

ATTORNEY/AGENT INFORMATION:

NAME: FENTON ESQ., GILLIAN M.

REGISTRATION NUMBER: 36,508

REFERENCE/DOCKET NUMBER: CRP-074DV

TELECOMMUNICATION INFORMATION:

TELEPHONE: (508) 435-9001

TELEFAX: (508) 435-6951

INFORMATION FOR SEQ. ID NO: /note- "MOP-2 (MATURE FORM)"

SEQUENCE CHARACTERISTICS:

LENGTH: 139 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

NAME/KEY: Protein

LOCATION: 1..139

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-10-050-050-8

Query Match 36.0%; Score 46.5; DB 15; Length 139;

Best Local Similarity 47.8%; Pred. No. 35;

Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 3 LKYOELSNVLPPrKMGlf 25

DB

5 LKRQPKTKNE-LPHPNKLPgIf 26

RESULT 11

US-08-957-425-27
Sequence 27, Application US/08957425
Publication No. US20030069401A1

GENERAL INFORMATION:

APPLICANT: OPPERMANN, HERMANN

OZKAYNAK, ENGIN

KUBERASAMPATH, THANGAVEL

ROEGER, DAVID C.

PANG, ROY H.L.

TITLE OF INVENTION: OSTEOGENIC DEVICES

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: 53 STATE STREET

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/957,425

FILING DATE: 24-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/447,570

FILING DATE: 21-FEB-1992

APPLICATION NUMBER: US 810,560

FILING DATE: 20-DEC-1991

APPLICATION NUMBER: US 827,052

FILING DATE: 28-JAN-1992

APPLICATION NUMBER: US 660,162

FILING DATE: 22-FEB-1991

APPLICATION NUMBER: US 621,988

FILING DATE: 04-DEC-1990

APPLICATION NUMBER: US 621,849

FILING DATE: 04-DEC-1990

APPLICATION NUMBER: US 616,374

FILING DATE: 21-NOV-1990

APPLICATION NUMBER: US 600,024

FILING DATE: 18-OCT-1990

APPLICATION NUMBER: US 599,543

FILING DATE: 18-OCT-1990

APPLICATION NUMBER: US 579,865

FILING DATE: 07-SEP-1990

APPLICATION NUMBER: US 569,920

FILING DATE: 20-AUG-1990

APPLICATION NUMBER: US 483,913

FILING DATE: 22-FEB-1990

APPLICATION NUMBER: US 422,613

FILING DATE: 17-OCT-1989

APPLICATION NUMBER: US 315,342

FILING DATE: 23-FEB-1988

APPLICATION NUMBER: US 232,630

FILING DATE: 15-AUG-1988

APPLICATION NUMBER: US 179,460

FILING DATE: 08-APR-1988

ATTORNEY/AGENT INFORMATION:

NAME: PITCHER, EDMUND R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-001CP6

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/248-7000

TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-08-957-425-27

Query Match 36.0%; Score 46.5; DB 8; Length 399;
Best Local Similarity 47.8%; Pred. No. 1.1e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKVAQELSNVEVLPPPRKKGLF 25
DB 265 LKRQPKKTNE-LPHPNKLPGLF 286

RESULT 12

US-08-260-675-23
Sequence 23, Application US/08260675
Publication No. US2003010493A1

GENERAL INFORMATION:

APPLICANT: RUEGER, DAVID C
APPLICANT: KUBERASAMPATH, THANGAVEL
APPLICANT: OPPERMAN, HERMANN
APPLICANT: OKAYNAK, ENGIN
APPLICANT: PANG, ROY HL
APPLICANT: COHEN, CHARLES M
TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
TITLE OF INVENTION: REPAIR
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESS: TESTA HURWITZ & THIBEAULT
STREET: 55 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/260,675
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/126,100
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/922,813
FILING DATE:
APPLICATION DATA:
APPLICATION NUMBER: US 07/667,274
FILING DATE: 11-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/752,764
FILING DATE: 30-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FITCHER ESQ, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-070
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-260-675-23

Query Match 36.0%; Score 46.5; DB 8; Length 399;
Best Local Similarity 47.8%; Pred. No. 1.1e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKVAQELSNVEVLPPPRKKGLF 25
DB 265 LKRQPKKTNE-LPHPNKLPGLF 286

RESULT 13

US-10-350-747-2
Sequence 2, Application US/10350747
Publication No. US20030153072A1

GENERAL INFORMATION:

APPLICANT: Hogan, Bridgid L.M.
TITLE OF INVENTION: Compositions and Methods of Making
Embryonic Stem Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Panitch Schwarze Jacobs & Nadel, P.C.
STREET: 1601 Market Street, 36th Floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103-2398

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/350,747
FILING DATE: 24-Jan-2003
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/808,346
FILING DATE: 28-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9823-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-567-2020
TELEFAX: 215-567-2991
TELEX: 831-494

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-350-747-2

Query Match 36.0%; Score 46.5; DB 12; Length 399;
Best Local Similarity 47.8%; Pred. No. 1.1e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKVAQELSNVEVLPPPRKKGLF 25
DB 265 LKRQPKKTNE-LPHPNKLPGLF 286

RESULT 14

US-10-122-026-8
Sequence 8, Application US/10122026
Publication No. US20030105004A1

GENERAL INFORMATION:

APPLICANT: JONES, WILLIAM K
TUCKER, RONALD F
RUEGER, DAVID C

OPPERMANN, HERMANN
OKAYNAK, ENGIN
KUBERASAMPATH, THANGAVEL
TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
OF MATTER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patent Administrator, Testa, Hurwitz &
Thibault, LLP
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/122,026
FILING DATE: 29-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/402,542
FILING DATE: 13-MAR-1995
APPLICATION NUMBER: US 08/940,510
FILING DATE: 31-MAR-1993
APPLICATION NUMBER: US 08/029,335
FILING DATE: 04-MAR-1993
APPLICATION NUMBER: US 07/971,091
FILING DATE: 03-NOV-1992
APPLICATION NUMBER: US 07/946,235
FILING DATE: 16-SEP-1992
APPLICATION NUMBER: US 07/938,336
FILING DATE: 28-AUG-1992
APPLICATION NUMBER: US 07/923,780
FILING DATE: 31-JUL-1992
APPLICATION NUMBER: US 07/752,857
FILING DATE: 30-AUG-1991
APPLICATION NUMBER: US 07/752,764
FILING DATE: 30-AUG-1991
APPLICATION NUMBER: US 07/667,274
FILING DATE: 11-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: CAMACHO, JENNIFER A.
REGISTRATION NUMBER: 43,526
REFERENCE/DOCKET NUMBER: STK-060CN
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-10-122-026-8
Query Match 36.0%; Score 46.5; DB 15; Length 399;
Best Local Similarity 47.8%; Pred. No. 1.1e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 3 LKTAQELSNVLPPIKMKGLF 25
DB 265 LKRRQPKTNE-LPHPKLPGIF 286
RESULT 15
US-10-050-050-23
Sequence 23, Application US/10050050
Publication No. US20030125230A1
GENERAL INFORMATION:
APPLICANT: COHEN, CHARLES M.
CHALETTE, MARC F.

KUBERASAMPATH, THANGAVEL
ROEGER, DAVID C.
OPPERMANN, HERMANN
PANG, ROY H.L.
OKAYNAK, ENGIN
SMART, JOHN E.
TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING
PROLIFERATION OF EPITHELIAL CELLS.
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
STREET: 45 SOUTH STREET
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 01748
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/050,050
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/461,113
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/445,882
FILING DATE: 22-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: FENTON Esq., GILLIAN M.
REGISTRATION NUMBER: 36,508
REFERENCE/DOCKET NUMBER: CRP-074DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 435-9001
TELEFAX: (508) 435-6951
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-050-050-23
Query Match 36.0%; Score 46.5; DB 15; Length 399;
Best Local Similarity 47.8%; Pred. No. 1.1e+02;
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
QY 3 LKTAQELSNVLPPIKMKGLF 25
DB 265 LKRRQPKTNE-LPHPKLPGIF 286
Search completed: October 1, 2003, 10:37:54
Job time : 31.1471 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 : Search time 54.4118 Seconds
(without alignments)
72.928 Million cell updates/sec

Title: US-09-171-432a-48
Perfect score: 129
Sequence: 1 WLNPKKINLADRMGLSGVOEIKQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: A.Geneseq.19Jun03.*
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*
25: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	18	AAW42969
2	129	100.0	25	25	ABG31728
3	129	100.0	2227	7	AAW42969
4	129	100.0	2227	11	AAW42969
5	129	100.0	2227	18	AAW42969
6	129	100.0	2227	21	AAW42969
7	129	100.0	2227	21	AAW42969
8	129	100.0	2227	21	AAW42969
9	129	100.0	2227	23	ABG31727

10	129	100.0	2227	23	ABG31728
11	129	100.0	2227	23	ABG31729
12	129	100.0	2227	23	ABG31730
13	129	100.0	2227	23	ABG31731
14	129	100.0	2227	24	ABG31732
15	129	100.0	2227	24	ABG31733
16	129	100.0	2227	24	ABG31734
17	129	100.0	2227	24	ABG31735
18	129	100.0	2227	24	ABG31736
19	129	100.0	2227	24	ABG31737
20	129	100.0	2227	24	ABG31738
21	129	100.0	2227	24	ABG31739
22	129	100.0	2227	24	ABG31740
23	129	100.0	2227	24	ABG31741
24	129	100.0	2227	24	ABG31742
25	129	100.0	2227	24	ABG31743
26	129	100.0	2227	24	ABG31744
27	129	100.0	2227	24	ABG31745
28	129	100.0	2227	24	ABG31746
29	129	100.0	2227	24	ABG31747
30	129	100.0	2227	24	ABG31748
31	129	100.0	2227	24	ABG31749
32	129	100.0	2227	24	ABG31750
33	129	100.0	2227	24	ABG31751
34	129	100.0	2227	24	ABG31752
35	129	100.0	2227	24	ABG31753
36	129	100.0	2227	24	ABG31754
37	129	100.0	2227	24	ABG31755
38	129	100.0	2227	24	ABG31756
39	129	100.0	2227	24	ABG31757
40	129	100.0	2227	24	ABG31758
41	129	100.0	2227	24	ABG31759
42	129	100.0	2227	24	ABG31760
43	129	100.0	2227	24	ABG31761
44	129	100.0	2227	24	ABG31762
45	129	100.0	2227	24	ABG31763

ALIGNMENTS

RESULT 1	AAW42969	standard; peptide: 25 AA.
ID	AAW42969	
AC	AAW42969	
XX		
DT	28-APR-1998	(first entry)
XX		
DE	Immunogenic Hepatitis A virus peptide YK-1757.	
XX		
KW	Immunogenic peptide; immunogenic epitope; P2A protein;	
XX		
OS	Synthetic.	
OS	Hepatitis A virus.	
XX		
PN	WO9740147-A1.	
XX		
PD	30-OCT-1997.	
XX		
PF	18-APR-1997.	97WO-US06891.
XX		
PR	19-APR-1996.	96US-0015644.
XX		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX		
PI	Fields HA, Khudayakov YE.	
XX		
DR	WPI, 1997-535831/49.	
XX		
PT	Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an	
PT	immune response to HAV in a mammal or to detect the presence of	

PT antibodies against HAV in a mammal
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC The present immunogenic peptide corresponds to an immunogenic
CC epitope of the Hepatitis A virus (HAV). The peptide is substantially
CC similar to a portion of the amino acid sequence of the P2A protein of HAV
CC corresponding to amino acids 792-980. Compositions containing the
CC peptide can be used to induce an immune response to HAV in a mammal.
CC The peptide can also be used to detect the presence of antibodies
CC against HAV in mammalian serum. The peptide can also be used to make an
CC antibody against HAV by administering the peptide to a mammal.
XX
SQ Sequence 25 AA:
Query Match 100.0%; Score 129; DB 18; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKKINLADRMGLSGVDEIKKQ 25
1 WLNPKKINLADRMGLSGVDEIKKQ 25
RESULT 2
AAB69448
ID AAB69448 standard; Peptide: 25 AA.
XX
AC AAB69448;
XX
DT 20-APR-2001 (first entry)
XX
DE Synthetic HAV P2A peptide, SEQ ID NO: 48.
XX
KW Hepatitis A virus; HAV; Immunogen; Immunostimulant; Virucide; Vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN W0200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US19267.
XX
PR 15-JUL-1999; 99US-0144412.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
Fields HA, Khudyakov YE;
XX
DR WPI: 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines -
XX
PS Claim 13; Page 99; 130pp; English.
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IgM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy
CC end of the peptides enhances the IgM antibody reactivity.
XX
SQ Sequence 25 AA:
Query Match 100.0%; Score 129; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKKINLADRMGLSGVDEIKKQ 25
1 WLNPKKINLADRMGLSGVDEIKKQ 25
Db
RESULT 3
AAP60066
ID AAP60066 standard; Protein: 2227 AA.
XX
AC AAP60066;
XX
DT 25-MAR-2003 (updated)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of viral I434 polypeptide encoded by the complete
DE nucleotide sequence of the HAV genome.
XX
KW Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
FH Key
FH Region Location/Qualifiers
FT 1..245
FT /label= P1.1A
FT 246..491
FT /label= 1B
FT 492..836
FT /label= 1C
FT 837..980
FT /label= P2.2A
FT 981..1076
FT /label= 2B
FT 1077..1422
FT /label= 2C
FT 1423..1484
FT /label= P3.3A
FT 1485..1507
FT /label= 3B
FT 1508..1678
FT /label= 3C
FT 1679..2227
FT /label= 3D
XX
PN EP199480-A.
XX
XX
PD 29-OCT-1986.
XX
PF 03-APR-1986; 86EP-0302465.
XX
PR 03-APR-1985; 85US-0719329.
XX
PA (CHIR) CHIRON CORP.
XX
PI Dlna D, Potter SJ, Vannest GA, Caput D;
XX WPI: 1986-286213/44.
XX N-PSDB; AAN60080.
XX
DR Hepatitis A virus nucleotide sequence and polypeptide - and use
DR in prodn. of vaccines and diagnostic probes
XX
PS Claim 5; Fig 1; 18pp; English.
XX
CC AAN60080 and oligonucleotide fragments are useful in detection of

XX (USSH) US SEC DEPT HEALTH.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Emerson SU, Purcell RH, Raychaudhuri G;
XX
DR WPI: 1997-535850/49.
DR N-PSDB: AAT93023.
PT Human attenuated HAV genome containing simian HAV 2C gene - useful
XX as vaccines against HAV infection
XX
PS Disclosure; Fig 13A-D; 66pp; English.
XX
CC This protein sequence is encoded by the human hepatitis A virus
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey
CC kidney cells. A claimed DNA construct (I) comprises a genome of
CC HAV, where the genome is a human attenuated HAV genome in which a
CC region of the 2C gene has been replaced by a corresponding region
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
CC region of the 2C gene from AGM-27 contained in the construct
CC preferably encodes amino acids 120-328 of the 2C protein, amino
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
CC transcript of (1); (2) a cell transfected with (1) or the RNA
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
CC its RNA transcript, can be used as a vaccine for preventing HAV in
CC a mammal. (1) or the RNA transcript can also be used to stimulate
CC the production of protective antibodies in the mammal.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3, 1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPRIINLADRMIGLSGVOEIKQ 25
DB 956 WLNPRIINLADRMIGLSGVOEIKQ 980
RESULT 6
AAB18607
ID AAB18607 standard; Protein; 2227 AA.
AC AAB18607;
XX
XX 15-JAN-2001 (first entry)
XX
XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX Hepatitis A virus.
XX OS US6113912-A.
XX PN
XX PD 05-SEP-2000.
XX
XX PE 07-JUN-1995; 95US-0475886.
XX
XX PR 18-SEP-1992; 92US-0947338.
XX PR 17-SEP-1993; 93WO-US08610.
XX PR 10-MAR-1995; 95US-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI: 2000-586464/55.
XX DR N-PSDB: AAA75476.
XX

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
XX
PS Disclosure; Fig 6A-K; 72pp; English.
XX
CC The present sequence is derived from a wild type hepatitis A virus
CC (HAV) strain HM-174. The sequence is modified to produce HAV which
CC are adapted to growth in the human fibroblast-like cell line MRC-5.
CC The HAV is able to propagate in MRC-5 cells and retain appropriate
CC attenuation. It is useful as a live vaccine for prophylaxis of
CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3, 1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WLNPRIINLADRMIGLSGVOEIKQ 25
DB 956 WLNPRIINLADRMIGLSGVOEIKQ 980
RESULT 7
AAB18608
ID AAB18608 standard; Protein; 2227 AA.
AC AAB18608;
XX
XX 15-JAN-2001 (first entry)
XX
XX
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX DE
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX P-35 virus.
XX KW
XX P-35 virus.
XX KM
XX
XX Hepatitis A virus.
XX OS
XX PN US6113912-A.
XX
XX PD 05-SEP-2000.
XX
XX PE 07-JUN-1995; 95US-0475886.
XX
XX PR 18-SEP-1992; 92US-0947338.
XX PR 17-SEP-1993; 93WO-US08610.
XX PR 10-MAR-1995; 95US-0397232.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI: 2000-586464/55.
XX DR N-PSDB: AAA75477.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type
XX
XX PS Disclosure; Columns 67-78; 72pp; English.
XX
XX CC The present sequence is derived from passage 35 of a wild type
XX CC hepatitis A virus (HAV) strain HM-174. The resulting virus is
XX CC designated P-35 virus. The sequence is modified to produce HAV which
XX CC are adapted to growth in the human fibroblast-like cell line MRC-5.
XX CC The HAV is able to propagate in MRC-5 cells and retain appropriate
XX CC attenuation. It is useful as a live vaccine for prophylaxis of
XX CC hepatitis A in humans and other primates.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 129; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.1e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 8

AAB18609
ID AAB18609 standard; Protein; 2227 AA.

AC AAB18609;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

HA; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
HAV 4380.

Hepatitis A virus.

US6113912-A.

05-SEP-2000.

07-JUN-1995; 95US-0475886.

18-SEP-1992; 92US-0947338.

17-SEP-1993; 93WO-US08610.

10-MAR-1995; 95US-0397232.

(USSR) US DEPT HEALTH & HUMAN SERVICES.

D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

WPI: 2000-586464/55.

N-PSDB: AAA75478.

Novel live hepatitis A virus adapted to growth in human fibroblast cell

line useful as vaccine for protecting humans against hepatitis A virus

infection, has modified genome compared to wild type

Disclosure; Columns 93-104; 72pp; English.

The present sequence is derived from a live attenuated hepatitis A

virus (HAV) of the invention, designated HAV 4380. The sequence is

produced by modifying wild type HAV strain HM-174. The HAV of the

invention are adapted to growth in the human fibroblast-like cell

line MRC-5. The HAV is able to propagate in MRC-5 cells and retain

appropriate attenuation. It is useful as a live vaccine for prophylaxis

of hepatitis A in humans and other primates.

Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.1e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25

DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 9

ABG31727

ID ABG31727 standard; Protein; 2227 AA.

AC ABG31727;

DT 29-NOV-2002 (first entry)

DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

HA; strain HM-175; MRC-5 cell; human fibroblast; virucide.

Hepatitis A virus strain HM-175.

US6423318-B1.

23-JUL-2002.

31-AUG-2000; 2000US-0653499.

07-JUN-1995; 95US-0475886.

17-SEP-1993; 93US-0397232.

17-SEP-1993; 93WO-US08610.

(USSR) US DEPT HEALTH & HUMAN SERVICES.

(SMK) SMITHKLINE BEECHAM BIOLOGICALS.

Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

WPI: 2002-680946/73.

N-PSDB: ABS2787.

New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

in MRC-5 cells, useful for preparing a vaccine against HAV infection

in MRC-5 cells, useful for preparing a vaccine against HAV infection

Disclosure; Fig 6; 71pp; English.

The invention relates to a polynucleotide which encodes a hepatitis A

virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

line). The polynucleotide is useful for preparing a vaccine against

hepatitis A virus infection. This sequence represents a hepatitis A virus

strain HM-175 polypeptide.

Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.1e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25

DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 10

ABG31728

ID ABG31728 standard; Protein; 2227 AA.

AC ABG31728;

DT 29-NOV-2002 (first entry)

Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.

Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

virucide; mutant; PHAV/7; muten.

Hepatitis A virus strain HM-175.

Synthetic.

Key Location/Qualifiers

FT MISC-difference 963

FT MISC-difference 764

FT MISC-difference 821

FT MISC-difference 1052

FT MISC-difference 1062

FT MISC-difference /note- "Wild-type Gly substituted by Ala"

```
FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"
FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"
FT
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX WPI: 2002-680946/73.
XX DR N-PSDB; ABS52788.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Example 3; Column 67-78; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents a hepatitis A virus
XX mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
XX Sequence 2227 AA:
XX
XX Query Match 100.0%; Score 129; DB 23; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 3,1e-12;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WLNPRTKINLADRMGLSGVQEIKEQ 25
XX I|||||
XX 956 WLNPRTKINLADRMGLSGVQEIKEQ 980
XX
XX RESULT 11
XX ABG31729
XX ID ABG31729 standard; Protein; 2227 AA.
XX
XX ABG31729;
XX
XX 29-NOV-2002 (first entry)
XX
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX HAV 4380.
XX
XX Hepatitis A virus strain HM-175.
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
```

```
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX WPI: 2002-680946/73.
XX DR N-PSDB; ABS52789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Disclosure; Column 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents an attenuated
XX hepatitis A virus 4830 polypeptide.
XX
XX Sequence 2227 AA:
XX
XX Query Match 100.0%; Score 129; DB 23; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 3,1e-12;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WLNPRTKINLADRMGLSGVQEIKEQ 25
XX I|||||
XX 956 WLNPRTKINLADRMGLSGVQEIKEQ 980
XX
XX RESULT 12
XX AAE19899
XX ID AAE19899 standard; Protein; 2227 AA.
XX
XX AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX
XX Hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX
XX WO200213855-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-1B01808.
XX
XX 17-AUG-2000; 2000US-225767P.
XX 29-AUG-2000; 2000US-228175P.
XX 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPER AB.
XX
XX Sallberg M, Hultgren C;
XX WPI: 2002-241837/29.
XX DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus
```


AC AB008641;
 XX
 DT 03-JUN-2003 (first entry)
 XX
 DE Attenuated hepatitis A virus (4380) strain HM-175.
 XX
 KW Hepatitis A virus; HAV, virucide; hepatotropic; antiinflammatory;
 XX vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
 OS Hepatitis A virus strain HM-175.
 XX
 PN US2002176869-A1.
 XX
 PD 28-NOV-2002.
 XX
 PF 29-APR-2002; 2002US-0135988.
 XX
 PR 07-JUN-1995; 95US-0475886.
 PR 31-AUG-2000; 2000US-0653499.
 PR 18-SEP-1992; 92US-0947338.
 PR 17-SEP-1993; 93WO-US08610.
 PR 17-APR-1995; 95US-0397232.
 XX
 PA (FUNK/) FUNKHOUSER A W.
 PA (EMER/) EMERSON S U.
 PA (FURC/) FURCELL R H.
 PA (DHON/) D'HONDT E.
 PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
 XX
 DR WPI: 2003-352605/33.
 DR N-PSDB; ABX93475.
 XX
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
 PT useful in vaccines for protecting primates against hepatitis infection
 PT and disease -
 XX
 PS Disclosure; Page 45-51; 70pp; English.
 XX
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
 CC a vaccine for protecting primates against hepatitis infection and
 CC disease. This is the amino acid sequence of an attenuated human
 CC hepatitis A virus (4380) strain HM-175.
 XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 24; Length 2227;
 Best Local Similarity 100.0%; Pred. NO.3.1e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKEDQ 25
 ||||||||||||||||||||
 Db 956 WLNPKKINLADRMGLSGVOEIKEDQ 980

Search completed: October 1, 2003, 09:56:47
 Job time : 55.418 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 16.0294 Seconds
(without alignments)
149.988 Million cell updates/sec

Title: US-09-171-432a-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2227	1 GNNYHM	genome polyprotein
2	129	100.0	2227	1 GNNYHR	genome polyprotein
3	129	100.0	2227	1 GNNYMK	genome polyprotein
4	129	100.0	2227	1 GNNYHB	genome polyprotein
5	129	100.0	2230	1 GNNYSA	genome polyprotein
6	51.5	39.9	656	2 B82056	glutathione-regula
7	50	38.8	172	2 S33416	heat shock protein
8	49	38.0	327	1 RDVZAS	ribonucleoside-dip
9	48	37.2	158	2 S64321	hypothetical prote
10	47.5	36.8	482	2 S27608	cysteine proteinas
11	47.5	36.8	653	2 G82971	probable ferredoxi
12	47	36.4	145	2 B25199	heat shock 16k pro
13	47	36.4	381	2 T18876	hypothetical prote
14	46.5	36.0	146	2 B70142	ribosomal protein
15	46	35.7	342	2 H81317	probable lipopolys
16	46	35.7	342	2 H86336	hypothetical prote
17	46	35.7	518	2 S72432	MG096 homolog D09
18	46	35.7	632	2 S73431	probable oxidoredu
19	46	35.7	671	2 AF0042	probable oxidoredu
20	46	35.7	877	2 S58824	amido transferase C
21	45.5	35.3	201	2 H81274	ceos protein Bur
22	45.5	35.3	1027	2 T43024	isoleucine-tRNA 11
23	45	34.9	1042	2 H70203	hypothetical prote
24	45	34.9	1061	2 H90084	hypothetical prote
25	44	34.1	152	2 S07505	endoexyribonucle
26	44	34.1	165	2 S73194	hypothetical prote
27	44	34.1	278	2 B83288	probable maltose o
28	44	34.1	370	2 H70423	oxygen-independent
29	44	34.1	408	2 AC2295	succinyl-CoA synth

30	44	34.1	412	2 C96816	hypothetical prote
31	44	34.1	414	2 T06303	enoyl-CoA hydratase
32	44	34.1	512	2 T48462	cytochrome P450-11
33	44	34.1	646	2 S72609	GTP-binding membra
34	44	34.1	653	2 G70683	probable lepa - My
35	44	34.1	677	2 B82870	DNA topoisomerase
36	44	34.1	788	2 S67595	hypothetical prote
37	44	34.1	994	2 S16739	env protein - siml
38	44	34.1	1770	2 S56221	hypothetical prote
39	43.5	33.7	388	1 S72995	alanine racemase (
40	43.5	33.7	395	2 AD0380	multidrug efflux p
41	43.5	33.7	482	2 A38533	transcription acti
42	43	33.3	145	2 B24289	heat shock protein
43	43	33.3	148	2 G69125	ribosomal protein
44	43	33.3	180	2 AC0056	conserved hypothet
45	43	33.3	207	2 C49905	protein secretion

ALIGNMENTS

RESULT 1

GNNYHM

N: genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

B: RNA-directed RNA polymerase (BC 2.7.7.48), protein 3D

C: Species: human hepatitis A virus

A: Note: host Homo sapiens (man)

C: Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C: Accession: A25981

R: Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A: Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A: Reference number: A25981; PMID:87061253; PMID:3023706

A: Accession: A25981

A: Molecule type: genomic RNA

A: Residues: 1-2227 <COH>

A: Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C: Superfamily: hepatitis A virus genome polyprotein

C: Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP2>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches .25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 2

GNNYHR

N: genome polyprotein - human hepatitis A virus

B: RNA polymerase (BC 2.7.7.48), protein 3D

C: Species: human hepatitis A virus

A: Note: host Homo sapiens (man)

C: Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C: Accession: A03903

R: Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A: Title: Primary structure and gene organization of human hepatitis A virus.

```

A:Reference number: A03903; MWID:85190549; PMID:2966127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAD>
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597
C:Superfamily: hepatitis A virus genome polypeptide
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match          100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.le-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVGEIKED 25
|||||
956 WLNPKKINLADRMGLSGVGEIKED 980

```

```

RESULT 3
GNNYMK
genome polyprotein - human hepatitis A virus (strain HM-175/TMK-5, attenuated HAV)
NA:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
A:Note: host Homo sapiens (man)
C:Species: human hepatitis A virus
C:Accession: A94149; A25014; A94508
C:Date: 30-Jun-1988 #sequence,revision 30-Jun-1988 #text_change 16-Jul-1999
R:Chen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feltnstone, S.M.; Purcell, F
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A>Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:g329594; PIDN:AAA45471.1; PID:g329595
A:Note: Submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:1-6401/Product: coat protein 1B #status predicted <P1B>
F:1-836-836/Product: coat protein 1C #status predicted <P1C>
F:1-837-980/Product: core protein 2A #status predicted <P2A>
F:1-981-1076/Product: core protein 2B #status predicted <P2B>
F:1-1077-1422/Product: core protein 2C #status predicted <P2C>
F:1-1423-1484/Product: protein 3A #status predicted <P3A>
F:1-1485-1507/Product: protein 3B #status predicted <P3B>
F:1-1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1-1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match          100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. NO. 1.1e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      1 WLNPKKINTLADRMGLSGVOEIKKEQ 25
        |||||||
-Db     956 WLNPKKINTLADRMGLSGVOEIKKEQ 980

genome polyprotein - human hepatitis A virus (strain MBH)
NA:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pr
VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

```

C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
S:Paul, A.V.; Tade, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhard
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso
A:Reference number: J50303; MUID:86045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1..2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: coat protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <PG>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

```

Best local similarity 100.0%; Pred. No. 1, le-10;
Matches 25; Conservatively 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  WLNPKKINLADRMIGLSGVOEIKQ 25
          |||||
Db       956  WLNPKKINLADRMIGLSGVOEIKQ 980

RESULT 5
GNNYSA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
Submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:g222357; PIDN:BA00766.1; PID:g2223598
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure
A:Reference number: J01080; M01D:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
Submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL>
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA033490.1; PID:g930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian
A:Reference number: S03965; M01D:89232160; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
C:1-217/Product: coat protein 1A #status predicted <C1A>

```


F:28-249/Product: coat protein 1B #status predicted <ClB>
 F:250-495/Product: coat protein 1C #status predicted <ClC>
 F:496-795/Product: coat protein 1D #status predicted <ClD>
 F:796-984/Product: core protein 2A #status predicted <C2A>
 F:985-1091/Product: core protein 2B #status predicted <C2B>
 F:1092-1426/Product: core protein 2C #status predicted <C2C>
 F:1427-1498/Product: protein 3A #status predicted <P3A>
 F:1499-1521/Product: protein 3B #status predicted <P3B>
 F:1522-1741/Product: protein 3C #status predicted <P3C>
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 129; DB 1; Length 2230;
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLPKKNLADRMGLSGVOEIKQ 25
 Db 960 WLPKKNLADRMGLSGVOEIKQ 984

RESULT 6

glutathione-regulated potassium-efflux system protein KeftB VC2606 [Imported] - Vibrio cholerae
 C:Species: Vibrio cholerae
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: B82056
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qian, H.; Dragol, I.; Sellers, R. L. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MOID:20406853; PMID:10952301
 A:Accession: B82056
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-656 <HEI>
 A:Cross-references: GB:AE004327; GB:AE003852; NID:g9657185; PIDN:AAF95747.1; GSPDB:GN001
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC2606
 A:Map position: 1
 C:Superfamily: glutathione-regulated potassium efflux system protein keftB

Query Match 39.9%; Score 51.5; DB 2; Length 656;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

2 LNPKKI-NLADRMGLSGVOEI 22
 133 LNPKKLMNLRGPIILGLGCAQV 154

RESULT 7

heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)
 C:Species: Nippostrongylus brasiliensis
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 13-Aug-1999
 C:Accession: S33416
 R:Tweedie, S.; Grigg, M.E.; Ingram, L.; Selkirk, M.E.
 A:Submitted to the EMBL Data Library, April 1993
 A:Description: The expression of a small heat shock homologue is developmentally regulated
 A:Reference number: S33416
 A:Accession: S33416
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-172 <TWB>
 A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866
 C:Superfamily: alpha-crystallin

Query Match 38.8%; Score 50; DB 2; Length 172;
 Best Local Similarity 43.5%; Pred. No. 5.2;
 Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 3 NPKKINLADRMGLSGVOEIKQ 25
 Db 77 NELKVLDDRDILVEGMQVKE 99

RESULT 8

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - African swine fever
 N:Alternate names: ribonucleotide reductase small chain
 C:Species: African swine fever virus, ASFV
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 06-Nov-1998
 C:Accession: B40568
 R:Boursnell, M.; Shaw, K.; Yanez, R.J.; Vinuela, E.; Dixon, L.
 A:Title: The sequences of the ribonucleotide reductase genes from African swine fever
 A:Reference number: A40568; MOID:91335775; PMID:1871976
 A:Accession: B40568
 A:Molecule type: DNA
 A:Residues: 1-327 <BOU>
 A:Cross-references: GB:M64728
 C:Function:
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside
 C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain
 C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; ox
 F:70,101,104,164,196,201/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #
 F:108/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 38.0%; Score 49; DB 1; Length 327;
 Best Local Similarity 31.8%; Pred. No. 14;
 Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

OY 1 WLPKKNLADRMGLSGVOEI 22
 Db 145 WMDPARNSLGERLYGFAAVEGI 166

RESULT 9

hypothetical protein YGR030C - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein G4068
 C:Species: Saccharomyces cerevisiae
 C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 19-Apr-2002
 C:Accession: S64321
 R:Rieger, M.; Mueller-Auer, S.; Bruckner, M.; Schaefer, M.
 A:Submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64071
 A:Accession: S64321
 A:Molecule type: DNA
 A:Residues: 1-158 <RIE>
 A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN000
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:POF6; MRPS:YGR030C
 A:Cross-references: SGD:S0003262
 A:Map position: 7R

Query Match 37.2%; Score 48; DB 2; Length 158;
 Best Local Similarity 66.7%; Pred. No. 9.4;
 Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KRINLADRMGLSGV 19
 Db 64 KOINMADRSLGLQOV 78

RESULT 10

cysteine proteinase tpr - Porphyromonas gingivalis
 C:Species: Porphyromonas gingivalis
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 04-Mar-2000
 C:Accession: S27608
 R:Boutreau, G.; Lapointe, H.; Ploquin, P.; Mayrand, D.
 A:Submitted to the EMBL Data Library, February 1992

A:Description: Cloning, expression and sequencing of a protease gene (tpr) from *Porphyromonas gingivalis*

A:Reference number: S27608

A:Accession: S27608

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-482 <BOO>

A:Cross-references: EMBL:M84471, NTD:G150851, PTD:AAA25652.1, PTD:G150852

A:Superfamily: Porphyromonas gingivalis cysteine proteinase tpr

	Query Match	Score	47.5;	DB 2,	length	482;
	Best Local Similarity	33.3%;	Pred.	No. 36;		
Matches	8; Conservative	9; Mismatches	6; Indels	1; Gaps	1;	
Oy	1 WLNPRKTIADRMGLSGVQETKE	24				
	: : : : : :					
Db	80 WSNPRTDIYERVLG-SSMODLSK	102				

RESULT 11
G82971
Probable ferredoxin PA5399 [Imported] - Pseudomonas aeruginosa (strain PA01)

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: G82971
R:Stover, C.K.; Plam, X.O.; Eryin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: G82971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-653 <STO>
A:Cross-references: GB:AE004952; GB:AE004091; NID:99951710; PIDN:AG08784.1; GSPDB:GN001
A:Experimental source: strain PA01
A:Genetics:
A:Gene: PA5399

	Query Match	36.8%	Score 47.5	DB 2	Length 653
Best Local Similarity	45.5%		Pred. No. 50		
Matches	10		Conservative 6	Mismatches 5	Indels 1
Gaps					1
QY	2	LNPKKINLADRMGLGSGVOEIK	23		
DB	264	LNPKKL-IQDMVIGSLAGNDAR	284		

199
 12
 C: shock 16K protein 2 - Caenorhabditis elegans
 C: Species: Caenorhabditis elegans
 C: Date: 16-Aug-1988 #sequence: revision 16-Aug-1988 #text: change 13-Aug-1999
 C: Accession: B25199
 J: Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.
 J: Biol. Chem. 261, 12006-12015, 1986
 A: Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis elegans
 A: Reference number: A92555; MUID:86304344; PMID:3017958
 A: Accession: B25199
 A: Molecule type: DNA
 A: Residues: 1-145 <JON>
 A: Cross-references: GB:M14334; NID:g156338; PIDN:AAA28071.1; PID:g156340
 A: Superfamily: alpha-crystallin

Query Match	36.4%	Score 47	DB 2	Length 145
Best Local Similarity	55.6%	Pred. NO. 12		
Matches	10	Conservative	2	Mismatches 6; Indels 0; Gaps 0;
QY	6	KINLADRLGISGVOEIK	23	
			:	
db	65	KINLIDGRILSIOGEOELK	82	

RESULT 13
T18876
hypothetical protein C03C10.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C:Accession: T18876
C:Authors: M.

Submitted to the EMBL Data Library, August 1994
A:Reference number: Z19036
A:Accession: T18876
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1381 <WIL>
A:Cross-references: EMBL:Z25677; PIDN:CSA84688.1; GSPDB:GN00021; CESP:C03C10.3
A:Experimental source: clone C03C10
A:Genetics:

C: Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain

Query Match	36.48;	Score 47;	DB 2;	Length 381;
Best Local Similarity	31.88;	Pred. No. 34;		
Matches	7;	Conservative	8;	Mismatches 7;
				Indels 0;
				Gaps 0;

```
QY      1 WLNPKKINLADRMIGLSGVOEI 22
        |::||:|:::|:|:|
Db      205 WISDKKASFAERLIAFAAVEGI 226
```

RESULT 14
B70142
ribosomal protein L13 (rpL1M) - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 14-Apr-2003
C:Accession: B70142
R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kierulff, A.R.; Quackenbush, J.; Salberg, S.; Hanson, M.; Vu
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403665
A:Accession: B70142
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-146 <RLE>
A:Cross-references: GB:AE001140; GB:AE000783; NID:92686233; PIDN:AMC66717.1; PID:g268
A:Experimental source: strain B31
A:Superfamily: ribosomal protein L13

36.0%;	Score 46.5;	DB 2;	Length 146;
Best Local Similarity	36.7%;	Pred. No. 15;	
Matches 11;	Conservative 4;	Mismatches 8;	Indels 7;
			Gaps 1;
QY	1 WLNPK-----INLADRMIGLSGVDEIK	23	
	::		
	::		
bb	11 WIKRTEVKKVVVIDADRLITGKVAIVYVK	40	

RESULT 15
H81317
probable lipopolysaccharide heptosyltransferase Cj1133 [imported] - Campylobacter jej
C/Species: Campylobacter jejuni
C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 30-Sep-2002
C/Accession: H81317
R/Perkins, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, J.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barr
Nature 403, 665-668, 2000
A/Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A/Reference number: AB1250; M01D:20150912; PMID:10686204
A/Accession: H81317
A/Status: preliminary

A: Molecule type: DNA
A: Residues: 1-342 <PAR>
A: Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB73388.1; PID:g696856
A: Experimental source: serotype O2, strain NCTC 11168
C: Genetics:
A: Gene: waac: Cj1133
C: Superfamily: ADP-heptose-LPS heptosyltransferase II

Query Match 35.7%; Score 46; DB 2; Length 342;
Best Local Similarity 42.9%; Pred. No. 42;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOE 21
| | | | | : | | | | |
DB 220 WGNVKEVEFAKEVLNLSGIDE 240

Search completed: October 1, 2003, 10:04:48
Job time : 18.0294 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 8.23529 Seconds

(without alignments)
142.760 Million cell updates/sec

Title: US-09-171-432a-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2226	1	P26580 hepatitis a
2	129	100.0	2226	1	P26581 hepatitis a
3	129	100.0	2226	1	P26582 hepatitis a
4	129	100.0	2227	1	P08617 hepatitis a
5	129	100.0	2227	1	P06441 hepatitis a
6	129	100.0	2227	1	P13901 hepatitis a
7	129	100.0	2230	1	P14553 simian hepa
8	50	38.8	172	1	Q07160 nipostrom
9	49	38.0	327	1	P26713 african swi
10	48	37.2	158	1	P53218 saccharomyc
11	48	37.2	334	1	P42492 african swi
12	48	37.2	622	1	O9rc9g streptomyc
13	47.5	36.8	482	1	P25806 porphyromon
14	47	36.4	145	1	P06582 caenorhabd
15	47	36.4	145	1	P42170 caenorhabd
16	46.5	36.0	146	1	O51314 caenorhabd
17	46	35.7	518	1	P75066 mycoplasma
18	46	35.7	632	1	P75066 mycoplasma
19	45.5	35.3	201	1	O9pny4 campylobact
20	45	34.9	1042	1	O51773 borrelia bu
21	44	34.1	152	1	P20314 bacterioph
22	44	34.1	165	1	P51370 porphyra pu
23	44	34.1	646	1	P53330 mycobacteri
24	44	34.1	653	1	P71739 mycobacteri
25	44	34.1	985	1	P23073 simian fo
26	44	34.1	1770	1	P43365 saccharomyc
27	43.5	33.7	285	1	O8x76 buchnera ap
28	43.5	33.7	388	1	P38056 mycobacteri
29	43.5	33.7	441	1	O8rfx6 fusobacteri
30	43.5	33.7	482	1	P29267 alcaligenes
31	43	33.3	145	1	P34696 caenorhabd
32	43	33.3	148	1	O26129 methanobact
33	43	33.3	180	1	O8zip1 yersinia pe

34	43	33.3	207	1	GSPJ_AERHY	P31739 aeromonas h
35	43	33.3	367	1	MX12_MOUSE	O08911 mus musculu
36	43	33.3	421	1	V162_METJA	O57626 methanococc
37	43	33.3	495	1	NSRA_HAEIN	P43915 haemophilus
38	43	33.3	810	1	SVRA_SYNY3	P74296 synecocyst
39	42.5	32.9	284	1	RP32_BUCAI	O05385 buchnera ap
40	42.5	32.9	602	1	LEPA_BRUME	O8ydb8 bruceella me
41	42.5	32.9	687	1	ILVA_YEAST	P07342 saccharomyc
42	42.5	32.9	1472	1	ATC9_YEAST	O12697 saccharomyc
43	42	32.6	143	1	HS16_CAEEL	P06581 caenorhabd
44	42	32.6	151	1	SOPC_HALRO	P81926 halocynthia
45	42	32.6	185	1	TD52_MOUSE	O62393 mus musculu

ALIGNMENTS

RESULT 1	ID	POLG_HPAV2	STANDARD:	PRT: 2226 AA.
AC	P26580;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].			
DE	Hepatitis A virus (strain 24a).			
OS	Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.			
OC	Hepatovirus.			
OX	NCBI_Taxid=12094;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-91162758; PubMed-1705995;			
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,			
RA	Cromeans T., Jansen R.W.;			
RT	Antigenic and genetic variation in cytopathic hepatitis A virus			
RT	variants arising during persistent infection: evidence for genetic			
RT	recombination.;			
RL	J. Virol. 65:2056-2065(1991).			
CC	-I- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +			
CC	[RNA](N).			
CC	-I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,			
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,			
CC	VP3, AND VP4.			
CC	-I- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	-I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL: M59810; AAA4568.1; -			
DR	MEROPS: C03.005; -			
DR	InterPro: IPR004004; Calic1_pol_hel.			
DR	InterPro: IPR000605; RNA_helicase.			
DR	InterPro: IPR007095; RNA_pol_DS_PS.			
DR	InterPro: IPR001205; RNA_pol_P3D.			
DR	InterPro: IPR007094; RNA_pol_PSV1r.			
DR	Pfam: PF00680; RNA_dep_RNA_pol.1.			
DR	Pfam: PF00910; RNA_helicase.1.			
DR	PRINTS: PR00918; CALICVIRUSNS.			
KW	RNA-directed RNA polymerase; Core protein; Transferrase;			
KW	Hydrolase; Coat protein; Thiol protease.			
FT	CHAIN 1 23			
FT	CHAIN 24 245			
FT	CHAIN 246 491			
FT	CHAIN 492 794			
FT	CHAIN 795 900			
FT	CHAIN CORE PROTEIN VP4 (P1A).			
FT	CHAIN COAT PROTEIN VP2 (P1B).			
FT	CHAIN COAT PROTEIN VP3 (P1C).			
FT	CHAIN COAT PROTEIN VP1 (P1D).			
FT	CHAIN CORE PROTEIN P2A.			

FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D684E2BF CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 2

POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID POLG_HPAV4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Felstone S.M., Crommons T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination.";
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M59809; AAA45469.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
CORE PROTEIN P2A.
CORE PROTEIN P2B.
CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251107 MW; 403B4C80809BE75 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 3

POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID POLG_HPAV8
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Felstone S.M., Crommons T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus variants arising during persistent infection: evidence for genetic recombination.";
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate + [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M59608; AAA45467.1; -.
DR PDB; 1QAT; 15-MAY-00.
DR MEROPS; C03.005; -.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
PROBABLE PROTEIN P3A.

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964463396C8DB CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
956 WLNPKKINLADRMGLSGVQEIKEQ 980
RESULT 4
POLG_HPAVLH STANDARD: PRT; 2227 AA.
AC POLG641: P06443; Q81082;
PT 01-AUG-1988 (rel. 08; Created)
01-AUG-1988 (rel. 08; Last sequence update)
15-SEP-2003 (rel. 42; Last annotation update)
Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain HM-175).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
Baroudy B.M.;
RT *Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.*;
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Attenuated;
RX MEDLINE=8715701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Deemer R.J., Feinstone S.M.,
Purcell R.H.;
RT *Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.*;
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Malzel J.V. Jr.,
Purcell R.H., Feinstone S.M.;
RT *Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.*;
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
RN [4]
RP CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
[RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTH: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC -----
CC EMBL: M14114; AAA5475.1; -
CC DR EMBL: M14707; AAA5465.1; -
CC DR EMBL: M14707; AAA5466.1; ALT_INIT.
CC DR EMBL: M16532; AAA5471.1; -
CC DR PIR: A25981; GNNYMK.
CC DR PIR: A94149; GNNYMK.
CC DR PDB: 1HAV; 23-DEC-96.
CC DR MEROPS: C03.005; -
CC DR InterPro: IPR004004; Calic1_pol_hel.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVir.
CC DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam: PF00910; RNA_helicase; 1.
CC DR PRINTS: PR009148; CALICVIRUSNS.
CC KW Polypeptide; Coat protein; Core protein; Transferrase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
CC FT CHAIN 1 23
CC FT CHAIN 24 245
CC FT CHAIN 246 491
CC FT CHAIN 492 836
CC FT CHAIN 837 980
CC FT CHAIN 981 1087
CC FT CHAIN 1088 1422
CC FT CHAIN 1423 1486
CC FT CHAIN 1497 1519
CC FT CHAIN 1520 1738
CC FT CHAIN 1739 2227
CC FT CHAIN 77 77
CC FT VARIANT 764 764
CC FT VARIANT 821 821
CC FT VARIANT 1052 1052
CC FT VARIANT 1062 1062
CC FT VARIANT 1118 1118
CC FT VARIANT 1151 1151
CC FT VARIANT 1163 1163
CC FT VARIANT 1277 1277
CC FT VARIANT 1500 1500
CC FT VARIANT 1805 1805
CC FT VARIANT 1930 1930
CC SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKKINLADRMGLSGVQEIKEQ 25
|||||
956 WLNPKKINLADRMGLSGVQEIKEQ 980
RESULT 5
POLG_HPAVLH STANDARD: PRT; 2227 AA.
AC POLG641:
PT 01-JAN-1988 (rel. 06; Created)
01-JAN-1988 (rel. 06; Last sequence update)
28-FEB-2003 (rel. 41; Last annotation update)
Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 1A).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
Merryweather J., van Nest G., Dina D.;

RT "Primary structure and gene organization of human hepatitis A virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02990; AAA45472.1; -.
DR PIR: A03903; GNNYHR.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calicel.pol.hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA.pol.DS.PS.
DR InterPro: IPR001205; RNA.pol.P3D.
DR InterPro: IPR007094; RNA.pol.PSVlr.
DR Pfam: PF00680; RNA_dep.RNA.pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINIS: PR00918; CALICIVIRUSNS.
KM Polypeptide: Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1076
FT CHAIN 1077 1422
FT CHAIN 1423 1484
FT CHAIN 1485 1507
FT CHAIN 1508 1678
FT CHAIN 1679 2227
SO SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKNINLADRMGLSGVOEIKQ 25
|||||
956 WLNPKNINLADRMGLSGVOEIKQ 980

RESULT 6
POLG_HPAVM STANDARD; PRT: 2227 AA.
ID POLG_HPAVM PRT: 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MBF).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E.,
RA Deinhardt F.;

RT "The entire nucleotide sequence of the genome of human hepatitis A
RT virus (isolate MBF).";
RL Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PPM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -
CC THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC ENTITIES REQUIRES A LICENSE AGREEMENT (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M20273; AAA45474.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA.pol.DS.PS.
DR InterPro: IPR001205; RNA.pol.P3D.
DR InterPro: IPR007094; RNA.pol.PSVlr.
DR Pfam: PF00680; RNA_dep.RNA.pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KM Polypeptide: Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SO SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 WLNPKNINLADRMGLSGVOEIKQ 25
|||||
956 WLNPKNINLADRMGLSGVOEIKQ 980

RESULT 7
POLG_HPAVS STANDARD; PRT: 2230 AA.
ID POLG_HPAVS PRT: 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGW-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balaayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGW-27: comparison of genome


```

RT structure and growth in cell culture with other HAV strains.*;
RL J. Gen. Virol. 72:1677-1683(1991).
(12)
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.*;
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00924; BAA00766.1; -
DR EMBL: X15461; CAA33490.1; -
DR PIR: A30470; GNNYSA.
DR MEROPS: C03.005; -
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR PolyProtein: Coat protein; Core protein; Transferase;
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
KW CHAIN 1 27
FT CHAIN 28 249 COAT PROTEIN VP4 (PIA).
FT CHAIN 250 495 COAT PROTEIN VP2 (PIB).
FT CHAIN 496 795 COAT PROTEIN VP3 (PIC).
FT CHAIN 796 984 COAT PROTEIN VP1 (PID).
FT CHAIN 985 1091 CORE PROTEIN P2A.
FT CHAIN 1092 1426 CORE PROTEIN P2B.
FT CHAIN 1427 1498 CORE PROTEIN P2C.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3A.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3B.
FT CHAIN 1742 2230 PROBABLE PROTEIN 3C.
FT CHAIN 2230 251296 RNA-DIRECTED POLYMERASE 3D.
FT CHAIN 251296 87B3230E324E1F19 CRC64;
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2230;
Best Local Similarity 100.0%; Pred. No. 1; 19e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKNLADRMGLSGVOEIKQ 25
DB 960 WLNPKNLADRMGLSGVOEIKQ 984

RESULT 8
HS20_NIPBR STANDARD; PRT; 172 AA.
AC 007160:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Heat shock protein homology (HSP20).
GN HSP20.
OS Nipostromylius brasiliensis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Strongylida;

```

```

OC Trichostrongyloidea; Heligmonellidae; Nipostromylinae;
OC Nipostromylius.
OX NCBI_TaxID=27635;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-94081860; PubMed-8259127;
RA Tweedle S., Grigg M.E., Ingram L., Selkirk M.E.;
RT The expression of a small heat shock protein homologue is
RT developmentally regulated in Nipostromylius brasiliensis.*;
RL Mol. Biochem. Parasitol. 61:149-154(1993).
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X71663; CAA50655.1; -
DR PIR: S33416; S33416.
DR InterPro: IPR001436; Crystallin_alpha.
DR InterPro: IPR002068; HSP20.
DR Pfam: PF00011; HSP20.1.
DR PRINTS: PR00299; ACRYSTALLIN.
DR PROSITE: PS01031; HSP20.1.
KW Heat shock.
SQ SEQUENCE 172 AA; 20227 MW; 2CDAA711CE60B1C0 CRC64;

Query Match 38.8%; Score 50; DB 1; Length 172;
Best Local Similarity 43.5%; Pred. No. 1.9;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 3 NPKKNLADRMGLSGVOEIKQ 25
DB 77 NPKKNLADRMGLSGVOEIKQ 99

RESULT 9
RIR2_ASFW2 STANDARD; PRT; 327 AA.
AC P26713:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE (Ribonucleoside-diphosphate reductase small chain (RC 1.17.4.1)
DE (Ribonucleoside-diphosphate reductase).
OS African swine fever virus (isolate Malawi I/1 20/1) (ASFV).
OS Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91335775; PubMed-1871976;
RA Boursnell M., Shaw K., Yanez R.J., Vinuela E., Dixon L.;
RT The sequences of the ribonucleoside reductase genes from African
RT swine fever virus show considerable homology with those of the
RT orthopoxvirus, vaccinia virus.*;
RL Virology 184:411-416(1991).
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodoxin + H(2)O -> ribonucleoside diphosphate + reduced
CC thiorodoxin.
CC -1- COFACTOR: BINDS 2 IRON IONS.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@sib-sib.ch).  
-----  
CC EMBL; M64728; -, NOT_ANNOTATED_CDS.  
DR PIR; B40568; RDVZAS.  
DR HSSP; P11157; IXSM.  
DR InterPro; IPR000358; Ribonucl_redtctse.  
DR Pfam; PF00268; ribonuc_red.sm.1.  
DR PROSITE; PS00368; RIBORED_SMALL; 1.  
KM Oxidoreductase; DNA replication; Iron.  
FT METAL 70 70 IRON 1 (BY SIMILARITY).  
FT METAL 101 101 IRON 1 AND 2 (BY SIMILARITY).  
FT METAL 104 104 IRON 1 (BY SIMILARITY).  
FT METAL 164 164 IRON 2 (BY SIMILARITY).  
FT METAL 198 198 IRON 2 (BY SIMILARITY).  
FT METAL 201 201 IRON 2 (BY SIMILARITY).  
FT ACT_SITE 108 108 BY SIMILARITY.  
SO SEQUENCE 327 AA; 38966 MW; E78508DBA1978F4B0 CRC64;  
  
Query Match 38.0%; Score 49; DB 1; Length 327;  
Best local Similarity 31.8%; Pred. No. 5.2;  
Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;  
  
OY 1 WLNPKKINADRMGLSGQGEI 22 ::::|::|:  
DB 145 WMDFARNSLGERTLVGFVAVEGI 166 ::||::|:  
  
RESULT 10  
POPE_YEAST STANDARD; PRY; 158 AA.  
ID POPE_YEAST  
AC P53218;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ribonucleases P/MRP protein subunit POPE (EC 3.1.26.5) (RNases P/MRP  
18.2 kDa subunit) (RNA processing protein POP6).  
GN POP6 OR YGR030C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
CX Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
NX NCBI_TaxId=4932;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=S288c;  
RX MEDLINE=97435481; PubMed=9290212;  
RISieger M., Brueckner M., Schaefer M., Mueller-Tuer S.;  
"Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
chromosome VII.";  
RL Yeast 13:1077-1090(1997).  
CC -I FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT  
GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.  
CC ALSO A COMPONENT OF RNASE MRP.  
CC -I CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-  
extra-nucleotide from tRNA precursor.  
CC -I SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P  
RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOTETY AND AT LEAST  
8 PROTEIN SUBUNITS; POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND  
PP1.  
CC -I SUBCELLULAR LOCATION: Nuclear (potential).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@sib-sib.ch).
```

```

DR SGD: S0003262; POP6.
DR GO: GO:0000172; C:ribonuclease mitochondrial RNA processing c. . .; IDA.
DR GO: GO:0005655; C:ribonuclease P complex; IDA.
DR GO: GO:0000171; F:ribonuclease MP activity; IDA.
DR GO: GO:0004526; F:ribonuclease P activity; IDA.
DR Hydrolase; Nucleolar protein; RNA processing; Coiled coil.
FT DOMAIN 51
FT COILED COIL (POTENTIAL)
SQ SEQUENCE 158 AA; 18210 MW; 6C27A73FAD521181 CRC64;

Query Match 37.28; Score 48; DB 1; Length 158;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 64 KOINMADRSGLQGV 78

RESULT 11
RIR2_ASFB7
ID RIR2_ASFB7 STANDARD; PRT; 334 AA.
PA2492;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleoside reductase).
GN F334L.
OS African swine fever virus (strain BA71V) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID-104498;
RN [1]
RA SEQUENCE FROM N.A.
RA yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RA Rodriguez J.F., Vinuela E.;
RT "Analysis of the complete nucleotide sequence of African swine fever
RT virus.";
RU Virology 208:249-278(1995).
CC -1 FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1 CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioedoxin.
CC -1 COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
CC -1 PATHWAY: DNA replication pathway; first step.
CC -1 SUBUNIT: Heterodimer of a large and a small chain.
CC -1 SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL; U18466; AAA5274.1; -.
DR HSSP; P1157; IAXM.
DR InterPro: IPR000358; Ribonuc_redctase.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DE Oxidoreductase; DNA replication; Iron.
FT METAL 77
FT METAL 77 IRON 1 (BY SIMILARITY).
FT METAL 108 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 111 IRON 1 (BY SIMILARITY).
FT METAL 171 IRON 2 (BY SIMILARITY).
FT METAL 205 IRON 2 (BY SIMILARITY).
FT METAL 208 IRON 2 (BY SIMILARITY).
FT ACT_SITE 115 IRON 2 (BY SIMILARITY).
FT ACT_SITE 115 BY SIMILARITY.
SQ SEQUENCE 334 AA; 39806 MW; 9501955E2FD719CF CRC64;

Query Match 37.28; Score 48; DB 1; Length 334;
Best Local Similarity 31.8%; Pred. No. 7.5;

```

Thu Oct 2 09:14:04 2003

Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIK 22

DB 152 WMDPRTTLGERTVGFVAVEGI 173

RESULT 12

LEPA_STRCO STANDARD: PRT; 622 AA.

AC 09RDC9;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE GTP-binding protein lepa.

GN LEPA OR SC02562 OR SCC77.29C.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinobacteriales; Streptomyces; Streptomycetaceae; Streptomyces.

NCBI_TaxID=1902;

[1]

SEQUENCE FROM N.A.

RC STRAIN-A3(2) / M145;

RX MEDLINE-21996410; PubMed-12000953;

RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,

RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,

RA Warren T., Watzorek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT *Complete genome sequence of the model actinomycete Streptomyces

RT coelicolor A3(2).*

RL Nature 417:141-147(2002).

CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.

LEPA_SBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

EMBL: AL939113; CAB66240.1; -.

HSRP: P13551; ZEPG.

HA MAP: MF_00071; -; 1.

DR InterPro: IPR000795; EF_GTPbind.

DR InterPro: IPR000640; EFG_C.

DR InterPro: IPR004161; EFTU_D2.

DR InterPro: IPR006297; Lepa.

DR InterPro: IPR005225; Small_GTP.

DR Pfam: PF00679; EFG_C.1.

DR Pfam: PF00009; GTP_EFTU.1.

DR Pfam: PF03144; GTP_EFTU_D2.1.

DR PRINTS: PR00315; ELONGATNFC.

DR TIGRFAMs: TIGR01393; lepa.1.

DR TIGRFAMs: TIGR00231; small_GTP.1.

DR PROSITE: PS00301; EFATOR_GTP.1.

KW GTP-binding; Complete proteome.

FT NP_BIND 26 33 GTP (BY SIMILARITY).

FT NP_BIND 94 98 GTP (BY SIMILARITY).

FT NP_BIND 148 151 GTP (BY SIMILARITY).

SO SEQUENCE 622 AA; 68378 MW; 835C76FA2A80C7C CRC64;

Query Match 37.2%; Score 48; DB 1; Length 622;

Best Local Similarity 57.9%; Pred. No. 14;

Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 KINLADRMGLSGVOEIK 24

DB 32 KSTLADRMGLTGVGEORQ 50

RESULT 13

TPR_PORGI STANDARD: PRT; 482 AA.

ID TPR_PORGI

AC P25806;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Thiol protease (EC 3.4.22.-).

GN TPR.

OS Porphyromonas gingivalis (Bacteroides gingivalis).

OC Bacteria; Bacteroidetes; Bacteroidales; Bacteroidales; Porphyromonadaceae; Porphyromonas.

NCBI_TaxID=837;

[1]

SEQUENCE FROM N.A.

RP STRAIN-W83;

RX MEDLINE-92347990; PubMed-1322368;

RA Bourgeois G., Lapointe H., Pelouquin P., Mayrand D.;

RT *Cloning, expression, and sequencing of a protease gene (tpr) from

RT Porphyromonas gingivalis W83 in Escherichia coli.*;

RL Infect. Immun. 60:3186-3192(1992).

CC -1- FUNCTION: THIOL PROTEASE. PROBABLY AN IMPORTANT VIRULENCE FACTOR.

CC -1- ENZYME REGULATION: INACTIVE BELOW 20 DEGREES CELSIUS AND PH 6.0.

CC INHIBITED BY DIVALENT CATIONS.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

EMBL: M84471; AAA25652.1; -.

DR PIR: S27608; S27608.

DR MEROPS: C01.090; -.

DR InterPro: IPR001300; Protease_C2.

DR InterPro: IPR000169; SHProl_acsite.

DR Pfam: PF00648; Peptidase_C2.1.

DR SMART: SM00230; Cyspc.1.

DR PROSITE: PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.

DR PROSITE: PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.

DR PROSITE: PS00139; THIOL_PROTEASE_CYS; 1.

KW Hydrolyase; Thiol protease; Virulence.

FT ACT_SITE 229 229 BY SIMILARITY.

FT ACT_SITE 407 407 BY SIMILARITY.

FT ACT_SITE 427 427 BY SIMILARITY.

SO SEQUENCE 482 AA; 55104 MW; 7E1E0786C2A70A50 CRC64;

Query Match 36.8%; Score 47.5; DB 1; Length 482;

Best Local Similarity 33.3%; Pred. No. 13;

Matches 8; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 WLNPKKINLADRMGLSGVOEIK 24

DB 80 WSNPRTTIVRYLG-SSMODLSK 102

RESULT 14

HS12_CAEEL STANDARD: PRT; 145 AA.

ID HS12_CAEEL

AC P06562;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Heat shock protein HSP16-2.

GN HSP16-2 OR Y46H3A.3.

OS Caenorhabditis elegans.

```

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86304344; PubMed=3017958;
RA Jones D., Rusanak R.H., Kay R.J., Candido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
RL Caenorhabditis elegans that is flanked by repetitive elements.";
RN J Biol. Chem. 261:12006-12015(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Stoneking T., Wohlmann P., Lennox S.;
RN Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RP REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
    FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14334; AAA28071.1; -.
DR EMBL; AC006774; AAF60615.1; -.
DR PIR; B25199; B25199
DR WormPep; Y46H3A.3; CE22002.
DR InterPro; IPR002068; HSP20.
DR Pfam; PF00011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
SQ SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;
OY 6 KINLADRLGSGVOEIK 23
DB 65 KINLGDRTLSTGGEOEIK 82
PRT;
ID 15
ID 2_CAEEL
AC P42170;
STANDARD; PRT; 381 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
DE RNR-2 OR C03C10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -1 FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1 CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioedoxin.
CC -1 COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).

```

```

CC -1- PATHWAY: DNA replication pathway; first step
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z35637; CAAB4688.1; -.
DR PIR; T18876; T18876.
DR HSSP; P11157; 1XSM.
DR WormPep; C03C10.3; CE00874.
DR InterPro; IPR000358; RibonucL_redctase.
DR Pfam; PF00268; ribonuc_red_smf.1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
DR KW Oxidoreductase; DNA replication; Iron.
FT METAL 130 130 IRON 1 (BY SIMILARITY).
FT METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 164 164 IRON 1 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT METAL 258 258 IRON 2 (BY SIMILARITY).
FT METAL 261 261 IRON 2 (BY SIMILARITY).
FT ACT_SITE 168 168 BY SIMILARITY.
SQ SEQUENCE 381 AA; 44289 MW; 75497147ABF36C59 CMC64;

Query Match 36.4%; Score 47; DB 1; Length 381;
Best Local Similarity 31.8%; Pred. No. 12;
Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

1 WLNRKTIADRMGLSGVQRI 22
1::|::|::|::|::|::|
205 WISKDKASFARLTAFAAVEI 226

```

Search completed: October 1, 2003, 09:57:57
Job time : 10.2353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 41.4706 Seconds
(without alignments)
155.564 Million cell updates/sec

Title: US-09-171-432a-48
Perfect score: 129
Sequence: 1 WLNPKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 830525 seqs, 258052604 residues

al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP_ARCHAEA:*
2: SP_BACTERIA:*
3: SP_FUNGI:*
4: SP_HUMAN:*
5: SP_INVERTEBRATE:*
6: SP_MAMMAL:*
7: SP_MNC:*
8: SP_ORGANELLE:*
9: SP_PHAGE:*
10: SP_PLANT:*
11: SP_RODENT:*
12: SP_VIRUS:*
13: SP_VERTEBRATE:*
14: SP_UNCLASSIFIED:*
15: SP_RVIRUS:*
16: SP_BACTERIAP:*
17: SP_ARCHAEP:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	251	12	O9ENP8
2	129	100.0	251	12	O9ENP9
3	129	100.0	251	12	O9ENQ4
4	129	100.0	251	12	O9ENQ2
5	129	100.0	251	12	O9ENN2
6	129	100.0	251	12	O9ENP2
7	129	100.0	251	12	O9ENQ6
8	129	100.0	251	12	O9ENP1
9	129	100.0	251	12	O9ENP4
10	129	100.0	251	12	O9ENQ9
11	129	100.0	251	12	O9ENN4
12	129	100.0	251	12	O9ENP5
13	129	100.0	251	12	O9ENQ3
14	129	100.0	251	12	O9ENP7
15	129	100.0	251	12	O9ENQ5
16	129	100.0	251	12	O9ENQ5

17	129	100.0	251	12	O9ENP9	O9ENP9 hepatitis a
18	129	100.0	251	12	O9ENP6	O9ENP6 hepatitis a
19	129	100.0	251	12	O9ENP6	O9ENP6 hepatitis a
20	129	100.0	251	12	O9ENQ8	O9ENQ8 hepatitis a
21	129	100.0	251	12	O9ENP3	O9ENP3 hepatitis a
22	129	100.0	251	12	O9ENQ0	O9ENQ0 hepatitis a
23	129	100.0	1124	12	O84780	O84780 hepatitis a
24	129	100.0	1161	12	O05794	O05794 hepatitis a
25	129	100.0	2216	12	O9WMA2	O9WMA2 hepatitis a
26	129	100.0	2218	12	O67824	O67824 hepatitis a
27	129	100.0	2218	12	O67817	O67817 hepatitis a
28	129	100.0	2225	12	O9DLJ2	O9DLJ2 hepatitis a
29	129	100.0	2227	12	O9WMA0	O9WMA0 hepatitis a
30	129	100.0	2227	12	O9WMA3	O9WMA3 hepatitis a
31	129	100.0	2227	12	O67825	O67825 hepatitis a
32	129	100.0	2227	12	O9WMA9	O9WMA9 hepatitis a
33	129	100.0	2227	12	O8QV03	O8QV03 hepatitis a
34	129	100.0	2227	12	O9WMA1	O9WMA1 hepatitis a
35	129	100.0	2227	12	O67826	O67826 hepatitis a
36	129	100.0	2227	12	O8V0N6	O8V0N6 hepatitis a
37	129	100.0	2227	12	O9IFH5	O9IFH5 hepatitis a
38	129	100.0	2227	12	O9WMA4	O9WMA4 hepatitis a
39	129	96.9	251	12	O9ENP0	O9ENP0 hepatitis a
40	125	96.9	251	12	O9ENP5	O9ENP5 hepatitis a
41	125	96.9	251	12	O9ENP7	O9ENP7 hepatitis a
42	124	96.1	2225	12	O9DWR1	O9DWR1 hepatitis a
43	124	96.1	2227	12	O8OR16	O8OR16 hepatitis a
44	121	93.8	251	12	O9ENP0	O9ENP0 hepatitis a
45	121	93.8	251	12	O9ENP8	O9ENP8 hepatitis a

ALIGNMENTS

RESULT 1
ID O9ENP8 PRELIMINARY; PRT; 251 AA.
AC O9ENP8;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A304;
RA Fujiwara K.;
RT "hepatitis A virus."
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB047665; BMB12173.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKINLADRLGLSGVQEIKEQ 25
DB 120 WLNPKINLADRLGLSGVQEIKEQ 144

RESULT 2

O9ENN9 PRELIMINARY; PRT; 251 AA.
AC O9ENN9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A68;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047674; BAB12182.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;
Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
ID Q9ENQ4 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ4;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A201;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047659; BAB12167.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;
Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
ID Q9ENQ2 PRELIMINARY; PRT; 251 AA.
AC Q9ENQ2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A205;
RA Fujiwara K.;

RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047661; BAB12169.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;
Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
ID Q9ENN2 PRELIMINARY; PRT; 251 AA.
AC Q9ENN2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A9;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047681; BAB12189.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;
Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
ID Q9ENP2 PRELIMINARY; PRT; 251 AA.
AC Q9ENP2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage: Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A5;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047671; BAB12179.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;
Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

RESULT 7

OGEN06 PRELIMINARY; PRT; 251 AA.

AC OGEN06: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RA Fujivara K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047657; BAB12165.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

RESULT 8

OGENR1 PRELIMINARY; PRT; 251 AA.

AC OGENR1: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RA Fujivara K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047652; BAB12160.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

RESULT 9

OGENP1

ID OGENP1 PRELIMINARY; PRT; 251 AA.

AC OGENP1: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RA Fujivara K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047672; BAB12180.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28663 MW; C7EA6FBD19A1619 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

RESULT 10

OGENP4 PRELIMINARY; PRT; 251 AA.

AC OGENP4: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RA Fujivara K.;
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047669; BAB12177.1; -.
FT NON_TER 1 251
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

RESULT 11

OGEN09 PRELIMINARY; PRT; 251 AA.

AC OGEN09: 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A159;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047654; BAB12162.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
|||||
SOUT 12
Q9ENM4 PRELIMINARY; PRT; 251 AA.
ID O9ENM4
AC O9ENM4
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A77;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047679; BAB12187.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKQ 25
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
|||||
RESULT 13
Q9ENP5 PRELIMINARY; PRT; 251 AA.
ID O9ENP5
AC O9ENP5
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A407;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER 1
FT NON_TER 251

FT NON_TER 251
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
|||||
RESULT 14
Q9ENQ3 PRELIMINARY; PRT; 251 AA.
ID O9ENQ3
AC O9ENQ3
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A204;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047660; BAB12168.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144
|||||
RESULT 15
Q9ENP7 PRELIMINARY; PRT; 251 AA.
ID O9ENP7
AC O9ENP7
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A306;
RA Fujiwara K.;
RT "Hepatitis A virus."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKQ 25
|||||

Db 120 WLNPKKINLADRMGLSGVQETKEQ 144

Search completed: October 1, 2003, 10:02:41
Job time : 41.4706 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 15.8824 Seconds
(without alignments)
66,600 Million cell updates/sec

Title: US-09-171-432a-48
Perfect score: 129
Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2227	3	US-08-475-886-2
2	129	100.0	2227	3	US-08-475-886-4
3	129	100.0	2227	3	US-08-475-886-6
4	129	100.0	2227	3	US-08-397-232-2
5	129	100.0	2227	3	US-08-397-232-4
6	129	100.0	2227	3	US-09-171-387-2
7	129	100.0	2227	4	US-09-653-499-2
8	129	100.0	2227	4	US-09-653-499-4
9	129	100.0	2227	4	US-09-653-499-6
10	49	38.0	229	3	US-09-248-528-17
11	49	38.0	229	3	US-09-549-108-17
12	49	38.0	229	3	US-09-549-111-17
13	49	38.0	229	3	US-09-549-106-17
14	49	38.0	229	3	US-09-550-394-17
15	47.5	36.8	699	4	US-09-252-991A-17077
16	46	33.7	877	4	US-09-165-396-5
17	45	34.9	373	4	US-09-134-001C-4798
18	44	34.1	495	4	US-09-107-532A-6687
19	42.5	32.9	393	4	US-09-198-452A-565
20	42	32.6	183	4	US-09-107-532A-5935
21	42	32.6	185	2	US-08-691-814B-12
22	42	32.6	185	4	US-09-250-609-11
23	42	32.6	185	4	US-09-250-611-11
24	42	32.6	271	4	US-09-252-991A-21522
25	42	32.6	343	4	US-09-252-991A-26240
26	42	32.6	514	3	US-09-385-028-5
27	42	32.6	514	4	US-09-726-614-5

28	42	32.6	638	2	US-08-846-762-95	Sequence 95, Appl
29	42	32.6	1250	1	US-08-441-139-9	Sequence 9, Appl
30	42	32.6	1404	4	US-08-801-308-1	Sequence 1, Appl
31	41	31.8	166	4	US-09-732-210-1606	Sequence 1606, Ap
32	41	31.8	184	2	US-08-715-204-5	Sequence 5, Appl
33	41	31.8	184	2	US-08-691-814B-50	Sequence 50, Appl
34	41	31.8	184	4	US-09-162-597-5	Sequence 5, Appl
35	41	31.8	184	4	US-09-250-609-13	Sequence 13, Appl
36	41	31.8	184	4	US-09-250-611-13	Sequence 6, Appl
37	41	31.8	366	3	US-08-746-883-6	Sequence 13, Appl
38	41	31.8	370	2	US-08-837-593-7	Sequence 6, Appl
39	41	31.8	416	3	US-08-946-329A-17	Sequence 7, Appl
40	41	31.8	416	3	US-08-567-357A-17	Sequence 17, Appl
41	41	31.8	416	3	US-08-729-743A-17	Sequence 17, Appl
42	41	31.8	416	3	US-08-349-498-17	Sequence 17, Appl
43	41	31.8	416	4	US-09-325-256-25	Sequence 25, Appl
44	41	31.8	416	5	PCT-US95-15463-17	Sequence 17, Appl
45	41	31.8	416	5	PCT-US95-15923-17	Sequence 17, Appl

ALIGNMENTS

```

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980
RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232

```

EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912

GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT FILING DATE: 1995-06-07
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 4
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S1
CURRENT FILING DATE: 1995-08/397,232A
EARLIER FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227

TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110

GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S1
CURRENT FILING DATE: 1995-08/397,232A
EARLIER FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734

GENERAL INFORMATION:
APPLICANT: RAYCHAUDHURI, GOPA;
EMERSON, SUZANNE, U.;
PURCELL, ROBERT, H.
TITLE OF INVENTION: SIMIAN-HUMAN HAV
HAVING A CHIMERIC 2C PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US97/06506
 FILING DATE: 18-APR-1997
 APPLICATION NUMBER: US60/015,642
 FILING DATE: 19-APR-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: William S. Feller
 REGISTRATION NUMBER: 26,728
 REFERENCE/DOCKET NUMBER: 2026-42290S1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 758-4800
 TELEFAX: (212) 751-6849
 INFORMATION FOR SEQ ID NO: 2
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2227 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 09-171-387-2

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2,9e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKEQ 25
 Db 956 WLNPKKINLADRMGLSGVOEIKEQ 980

RESULT 7
 US-09-653-499-2
 ; Sequence 2, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
 US-09-653-499-2

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2,9e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKEQ 25
 Db 956 WLNPKKINLADRMGLSGVOEIKEQ 980

RESULT 8
 US-09-653-499-4
 ; Sequence 4, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H
 APPLICANT: D'HONDT, ERIC
 TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 FILE REFERENCE: 20264262US2
 CURRENT APPLICATION NUMBER: US/09/653,499
 CURRENT FILING DATE: 2000-08-31
 PRIOR APPLICATION NUMBER: 08/475,886
 PRIOR FILING DATE: 1999-08-09
 PRIOR APPLICATION NUMBER: 07/947,338
 PRIOR FILING DATE: 1992-09-18
 PRIOR APPLICATION NUMBER: 08/397,232
 PRIOR FILING DATE: 1995-03-10
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 2227
 TYPE: PRT
 ORGANISM: Attenuated HAV (Pass 35), strain HM-175
 US-09-653-499-4

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2,9e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKEQ 25
 Db 956 WLNPKKINLADRMGLSGVOEIKEQ 980

RESULT 9
 US-09-653-499-6
 ; Sequence 6, Application US/09653499
 ; Patent No. 6423318
 ; GENERAL INFORMATION:
 ; APPLICANT: FUNKHOUSER, ANN W
 ; APPLICANT: EMERSON, SUZANNE U
 ; APPLICANT: PURCELL, ROBERT H
 ; APPLICANT: D'HONDT, ERIC
 ; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
 ; FILE REFERENCE: 20264262US2
 ; CURRENT APPLICATION NUMBER: US/09/653,499
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 08/475,886
 ; PRIOR FILING DATE: 1999-08-09
 ; PRIOR APPLICATION NUMBER: 07/947,338
 ; PRIOR FILING DATE: 1992-09-18
 ; PRIOR APPLICATION NUMBER: 08/397,232
 ; PRIOR FILING DATE: 1995-03-10
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 2227
 ; TYPE: PRT
 ; ORGANISM: Attenuated (4380) HAV, strain HM-175
 US-09-653-499-6

Query Match 100.0%; Score 129; DB 4; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2,9e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKEQ 25
 Db 956 WLNPKKINLADRMGLSGVOEIKEQ 980

RESULT 10
 US-09-248-528-17
 ; Sequence 17, Application US/09248528C
 ; Patent No. 6153415
 ; GENERAL INFORMATION:
 ; APPLICANT: Oriel, Patrick J
 ; APPLICANT: Padmakumar, Rugmini
 ; APPLICANT: Kim, Sang H

;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
;; FILE REFERENCE: MSU 4.1-401
;; CURRENT APPLICATION NUMBER: US/09/248,528C
;; EARLIER FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-248-528-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLNPKKINLADRL--GLSGVOEI 22
| | | : : : | | | : | : |
113 WENPKLVYVEKALLEGLSPVREV 136

RESULT 11
US-09-549-108-17
;; Sequence 17, Application US/09549108
;; Patent No. 6214603
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; APPLICANT: Kim, Sang H
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
;; FILE REFERENCE: MSU 4.1-486
;; CURRENT APPLICATION NUMBER: US/09/549,108
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-108-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLNPKKINLADRL--GLSGVOEI 22
| | | : : : | | | : | : |
113 WENPKLVYVEKALLEGLSPVREV 136

RESULT 12
US-09-549-111-17
;; Sequence 17, Application US/09549111
;; Patent No. 6228633
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; APPLICANT: Kim, Sang H
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
;; FILE REFERENCE: MSU 4.1-489
;; CURRENT APPLICATION NUMBER: US/09/549,111
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-111-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLNPKKINLADRL--GLSGVOEI 22
| | | : : : | | | : | : |
113 WENPKLVYVEKALLEGLSPVREV 136

RESULT 13
US-09-549-106-17
;; Sequence 17, Application US/09549106
;; Patent No. 6242342
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; APPLICANT: Kim, Sang H
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
;; FILE REFERENCE: MSU 4.1-487
;; CURRENT APPLICATION NUMBER: US/09/549,106
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-106-17

Query Match 38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLNPKKINLADRL--GLSGVOEI 22
| | | : : : | | | : | : |
113 WENPKLVYVEKALLEGLSPVREV 136

RESULT 14
US-09-550-394-17
;; Sequence 17, Application US/09550394
;; Patent No. 6287828
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; APPLICANT: Kim, Sang H
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrite
;; FILE REFERENCE: MSU 4.1-488
;; CURRENT APPLICATION NUMBER: US/09/550,394
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17

LENGTH: 229
 TYPE: PRT
 ORGANISM: Bacillus smith11
 US-09-550-394-17

Query Match 38.0%; Score 49; DB 3; Length 229;
 Best Local Similarity 41.7%; Pred. No. 1.4;
 Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLNPKKINLADRM--GLSGVOEI 22
 | | | : : : | | | : | :
 Db 113 WENPKLVKVEKALLEGISPVREV 136

RESULT 15
 US-09-252-991A-17077
 Sequence 17077, Application US/09252991A
 Patent No. 6551795

GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 17077
 LENGTH: 699

TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-17077

Query Match 36.8%; Score 47.5; DB 4; Length 699;
 Best Local Similarity 45.5%; Pred. No. 1.0;
 Matches 10; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

OY 2 LNPKNINLADRMGLSGVOEIK 23
 | | | | : : : | | : | :
 Db 310 LNPKNL-IDDMVIGLAGGMDAK 330

Search completed: October 1, 2003, 10:06:35
 Job time : 16.8824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 30.1471 seconds
(without alignments)
131.201 Million cell updates/sec

Title: US-09-171-432a-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRLGLSGVOEIKRQ 25

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 587654 seqs, 158212981 residues

587654

al number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	352	15	US-10-272-459-45
2	129	100.0	980	15	US-10-272-459-41
3	129	100.0	2227	15	US-09-929-955-12
4	129	100.0	2227	14	US-10-104-966-12
5	129	100.0	2227	14	US-10-135-988-2
6	129	100.0	2227	14	US-10-135-988-4
7	129	100.0	2227	14	US-10-135-988-6
8	48.5	37.6	1987	15	US-10-133-382-6
9	48.5	37.6	2013	15	US-10-133-382-2
10	48.5	37.6	2014	15	US-10-133-382-8
11	48.5	37.6	2040	15	US-10-133-382-4
12	45	34.9	309	11	US-09-510-332-67
13	45	34.9	325	10	US-09-764-864-1535
14	45	34.9	357	9	US-09-938-803-16
15	45	34.9	371	10	US-09-764-864-1118

16	45	34.9	622	15	US-10-156-761-13093	Sequence 13093, A
17 <td>44<td>34.1<td>1770<td>10<td>US-09-801-368-298</td><td>Sequence 298, App</td></td></td></td></td>	44 <td>34.1<td>1770<td>10<td>US-09-801-368-298</td><td>Sequence 298, App</td></td></td></td>	34.1 <td>1770<td>10<td>US-09-801-368-298</td><td>Sequence 298, App</td></td></td>	1770 <td>10<td>US-09-801-368-298</td><td>Sequence 298, App</td></td>	10 <td>US-09-801-368-298</td> <td>Sequence 298, App</td>	US-09-801-368-298	Sequence 298, App
18 <td>43<td>33.3<td>289<td>15<td>US-10-286-264-110</td><td>Sequence 110, App</td></td></td></td></td>	43 <td>33.3<td>289<td>15<td>US-10-286-264-110</td><td>Sequence 110, App</td></td></td></td>	33.3 <td>289<td>15<td>US-10-286-264-110</td><td>Sequence 110, App</td></td></td>	289 <td>15<td>US-10-286-264-110</td><td>Sequence 110, App</td></td>	15 <td>US-10-286-264-110</td> <td>Sequence 110, App</td>	US-10-286-264-110	Sequence 110, App
19 <td>43<td>33.3<td>358<td>9<td>US-09-861-451A-50</td><td>Sequence 50, Appl</td></td></td></td></td>	43 <td>33.3<td>358<td>9<td>US-09-861-451A-50</td><td>Sequence 50, Appl</td></td></td></td>	33.3 <td>358<td>9<td>US-09-861-451A-50</td><td>Sequence 50, Appl</td></td></td>	358 <td>9<td>US-09-861-451A-50</td><td>Sequence 50, Appl</td></td>	9 <td>US-09-861-451A-50</td> <td>Sequence 50, Appl</td>	US-09-861-451A-50	Sequence 50, Appl
20 <td>43<td>33.3<td>426<td>15<td>US-10-156-761-9490</td><td>Sequence 9490, App</td></td></td></td></td>	43 <td>33.3<td>426<td>15<td>US-10-156-761-9490</td><td>Sequence 9490, App</td></td></td></td>	33.3 <td>426<td>15<td>US-10-156-761-9490</td><td>Sequence 9490, App</td></td></td>	426 <td>15<td>US-10-156-761-9490</td><td>Sequence 9490, App</td></td>	15 <td>US-10-156-761-9490</td> <td>Sequence 9490, App</td>	US-10-156-761-9490	Sequence 9490, App
21 <td>43<td>33.3<td>495<td>9<td>US-09-815-242-11215</td><td>Sequence 11215, A</td></td></td></td></td>	43 <td>33.3<td>495<td>9<td>US-09-815-242-11215</td><td>Sequence 11215, A</td></td></td></td>	33.3 <td>495<td>9<td>US-09-815-242-11215</td><td>Sequence 11215, A</td></td></td>	495 <td>9<td>US-09-815-242-11215</td><td>Sequence 11215, A</td></td>	9 <td>US-09-815-242-11215</td> <td>Sequence 11215, A</td>	US-09-815-242-11215	Sequence 11215, A
22 <td>42<td>32.6<td>166<td>9<td>US-09-815-242-4942</td><td>Sequence 4942, App</td></td></td></td></td>	42 <td>32.6<td>166<td>9<td>US-09-815-242-4942</td><td>Sequence 4942, App</td></td></td></td>	32.6 <td>166<td>9<td>US-09-815-242-4942</td><td>Sequence 4942, App</td></td></td>	166 <td>9<td>US-09-815-242-4942</td><td>Sequence 4942, App</td></td>	9 <td>US-09-815-242-4942</td> <td>Sequence 4942, App</td>	US-09-815-242-4942	Sequence 4942, App
23 <td>42<td>32.6<td>166<td>9<td>US-09-815-242-10531</td><td>Sequence 10531, A</td></td></td></td></td>	42 <td>32.6<td>166<td>9<td>US-09-815-242-10531</td><td>Sequence 10531, A</td></td></td></td>	32.6 <td>166<td>9<td>US-09-815-242-10531</td><td>Sequence 10531, A</td></td></td>	166 <td>9<td>US-09-815-242-10531</td><td>Sequence 10531, A</td></td>	9 <td>US-09-815-242-10531</td> <td>Sequence 10531, A</td>	US-09-815-242-10531	Sequence 10531, A
24 <td>42<td>32.6<td>185<td>10<td>US-09-250-611-11</td><td>Sequence 11, Appl</td></td></td></td></td>	42 <td>32.6<td>185<td>10<td>US-09-250-611-11</td><td>Sequence 11, Appl</td></td></td></td>	32.6 <td>185<td>10<td>US-09-250-611-11</td><td>Sequence 11, Appl</td></td></td>	185 <td>10<td>US-09-250-611-11</td><td>Sequence 11, Appl</td></td>	10 <td>US-09-250-611-11</td> <td>Sequence 11, Appl</td>	US-09-250-611-11	Sequence 11, Appl
25 <td>42<td>32.6<td>299<td>9<td>US-09-825-882-16</td><td>Sequence 16, Appl</td></td></td></td></td>	42 <td>32.6<td>299<td>9<td>US-09-825-882-16</td><td>Sequence 16, Appl</td></td></td></td>	32.6 <td>299<td>9<td>US-09-825-882-16</td><td>Sequence 16, Appl</td></td></td>	299 <td>9<td>US-09-825-882-16</td><td>Sequence 16, Appl</td></td>	9 <td>US-09-825-882-16</td> <td>Sequence 16, Appl</td>	US-09-825-882-16	Sequence 16, Appl
26 <td>42<td>32.6<td>299<td>12<td>US-10-017-161-1832</td><td>Sequence 1832, App</td></td></td></td></td>	42 <td>32.6<td>299<td>12<td>US-10-017-161-1832</td><td>Sequence 1832, App</td></td></td></td>	32.6 <td>299<td>12<td>US-10-017-161-1832</td><td>Sequence 1832, App</td></td></td>	299 <td>12<td>US-10-017-161-1832</td><td>Sequence 1832, App</td></td>	12 <td>US-10-017-161-1832</td> <td>Sequence 1832, App</td>	US-10-017-161-1832	Sequence 1832, App
27 <td>42<td>32.6<td>405<td>14<td>US-10-029-180-78</td><td>Sequence 78, Appl</td></td></td></td></td>	42 <td>32.6<td>405<td>14<td>US-10-029-180-78</td><td>Sequence 78, Appl</td></td></td></td>	32.6 <td>405<td>14<td>US-10-029-180-78</td><td>Sequence 78, Appl</td></td></td>	405 <td>14<td>US-10-029-180-78</td><td>Sequence 78, Appl</td></td>	14 <td>US-10-029-180-78</td> <td>Sequence 78, Appl</td>	US-10-029-180-78	Sequence 78, Appl
28 <td>42<td>32.6<td>434<td>14<td>US-10-029-180-112</td><td>Sequence 112, App</td></td></td></td></td>	42 <td>32.6<td>434<td>14<td>US-10-029-180-112</td><td>Sequence 112, App</td></td></td></td>	32.6 <td>434<td>14<td>US-10-029-180-112</td><td>Sequence 112, App</td></td></td>	434 <td>14<td>US-10-029-180-112</td><td>Sequence 112, App</td></td>	14 <td>US-10-029-180-112</td> <td>Sequence 112, App</td>	US-10-029-180-112	Sequence 112, App
29 <td>42<td>32.6<td>608<td>15<td>US-10-153-668-196</td><td>Sequence 196, Appl</td></td></td></td></td>	42 <td>32.6<td>608<td>15<td>US-10-153-668-196</td><td>Sequence 196, Appl</td></td></td></td>	32.6 <td>608<td>15<td>US-10-153-668-196</td><td>Sequence 196, Appl</td></td></td>	608 <td>15<td>US-10-153-668-196</td><td>Sequence 196, Appl</td></td>	15 <td>US-10-153-668-196</td> <td>Sequence 196, Appl</td>	US-10-153-668-196	Sequence 196, Appl
30 <td>42<td>32.6<td>638<td>15<td>US-10-216-209-95</td><td>Sequence 95, Appl</td></td></td></td></td>	42 <td>32.6<td>638<td>15<td>US-10-216-209-95</td><td>Sequence 95, Appl</td></td></td></td>	32.6 <td>638<td>15<td>US-10-216-209-95</td><td>Sequence 95, Appl</td></td></td>	638 <td>15<td>US-10-216-209-95</td><td>Sequence 95, Appl</td></td>	15 <td>US-10-216-209-95</td> <td>Sequence 95, Appl</td>	US-10-216-209-95	Sequence 95, Appl
31 <td>42<td>32.6<td>1250<td>10<td>US-09-801-368-164</td><td>Sequence 364, App</td></td></td></td></td>	42 <td>32.6<td>1250<td>10<td>US-09-801-368-164</td><td>Sequence 364, App</td></td></td></td>	32.6 <td>1250<td>10<td>US-09-801-368-164</td><td>Sequence 364, App</td></td></td>	1250 <td>10<td>US-09-801-368-164</td><td>Sequence 364, App</td></td>	10 <td>US-09-801-368-164</td> <td>Sequence 364, App</td>	US-09-801-368-164	Sequence 364, App
32 <td>42<td>32.6<td>1404<td>9<td>US-09-811-045A-1</td><td>Sequence 1, Appl</td></td></td></td></td>	42 <td>32.6<td>1404<td>9<td>US-09-811-045A-1</td><td>Sequence 1, Appl</td></td></td></td>	32.6 <td>1404<td>9<td>US-09-811-045A-1</td><td>Sequence 1, Appl</td></td></td>	1404 <td>9<td>US-09-811-045A-1</td><td>Sequence 1, Appl</td></td>	9 <td>US-09-811-045A-1</td> <td>Sequence 1, Appl</td>	US-09-811-045A-1	Sequence 1, Appl
33 <td>42<td>32.6<td>2121<td>10<td>US-09-995-542-3</td><td>Sequence 3, Appl</td></td></td></td></td>	42 <td>32.6<td>2121<td>10<td>US-09-995-542-3</td><td>Sequence 3, Appl</td></td></td></td>	32.6 <td>2121<td>10<td>US-09-995-542-3</td><td>Sequence 3, Appl</td></td></td>	2121 <td>10<td>US-09-995-542-3</td><td>Sequence 3, Appl</td></td>	10 <td>US-09-995-542-3</td> <td>Sequence 3, Appl</td>	US-09-995-542-3	Sequence 3, Appl
34 <td>42<td>32.6<td>2167<td>10<td>US-09-995-542-2</td><td>Sequence 2, Appl</td></td></td></td></td>	42 <td>32.6<td>2167<td>10<td>US-09-995-542-2</td><td>Sequence 2, Appl</td></td></td></td>	32.6 <td>2167<td>10<td>US-09-995-542-2</td><td>Sequence 2, Appl</td></td></td>	2167 <td>10<td>US-09-995-542-2</td><td>Sequence 2, Appl</td></td>	10 <td>US-09-995-542-2</td> <td>Sequence 2, Appl</td>	US-09-995-542-2	Sequence 2, Appl
35 <td>42<td>32.6<td>5701<td>9<td>US-09-864-761-37319</td><td>Sequence 37319, A</td></td></td></td></td>	42 <td>32.6<td>5701<td>9<td>US-09-864-761-37319</td><td>Sequence 37319, A</td></td></td></td>	32.6 <td>5701<td>9<td>US-09-864-761-37319</td><td>Sequence 37319, A</td></td></td>	5701 <td>9<td>US-09-864-761-37319</td><td>Sequence 37319, A</td></td>	9 <td>US-09-864-761-37319</td> <td>Sequence 37319, A</td>	US-09-864-761-37319	Sequence 37319, A
36 <td>41<td>31.8<td>184<td>10<td>US-09-250-611-13</td><td>Sequence 13, Appl</td></td></td></td></td>	41 <td>31.8<td>184<td>10<td>US-09-250-611-13</td><td>Sequence 13, Appl</td></td></td></td>	31.8 <td>184<td>10<td>US-09-250-611-13</td><td>Sequence 13, Appl</td></td></td>	184 <td>10<td>US-09-250-611-13</td><td>Sequence 13, Appl</td></td>	10 <td>US-09-250-611-13</td> <td>Sequence 13, Appl</td>	US-09-250-611-13	Sequence 13, Appl
37 <td>41<td>31.8<td>184<td>15<td>US-10-205-823-417</td><td>Sequence 417, App</td></td></td></td></td>	41 <td>31.8<td>184<td>15<td>US-10-205-823-417</td><td>Sequence 417, App</td></td></td></td>	31.8 <td>184<td>15<td>US-10-205-823-417</td><td>Sequence 417, App</td></td></td>	184 <td>15<td>US-10-205-823-417</td><td>Sequence 417, App</td></td>	15 <td>US-10-205-823-417</td> <td>Sequence 417, App</td>	US-10-205-823-417	Sequence 417, App
38 <td>41<td>31.8<td>184<td>15<td>US-10-177-293-455</td><td>Sequence 455, App</td></td></td></td></td>	41 <td>31.8<td>184<td>15<td>US-10-177-293-455</td><td>Sequence 455, App</td></td></td></td>	31.8 <td>184<td>15<td>US-10-177-293-455</td><td>Sequence 455, App</td></td></td>	184 <td>15<td>US-10-177-293-455</td><td>Sequence 455, App</td></td>	15 <td>US-10-177-293-455</td> <td>Sequence 455, App</td>	US-10-177-293-455	Sequence 455, App
39 <td>41<td>31.8<td>197<td>15<td>US-10-146-473-74</td><td>Sequence 74, Appl</td></td></td></td></td>	41 <td>31.8<td>197<td>15<td>US-10-146-473-74</td><td>Sequence 74, Appl</td></td></td></td>	31.8 <td>197<td>15<td>US-10-146-473-74</td><td>Sequence 74, Appl</td></td></td>	197 <td>15<td>US-10-146-473-74</td><td>Sequence 74, Appl</td></td>	15 <td>US-10-146-473-74</td> <td>Sequence 74, Appl</td>	US-10-146-473-74	Sequence 74, Appl
40 <td>41<td>31.8<td>299<td>12<td>US-10-298-638-22</td><td>Sequence 22, Appl</td></td></td></td></td>	41 <td>31.8<td>299<td>12<td>US-10-298-638-22</td><td>Sequence 22, Appl</td></td></td></td>	31.8 <td>299<td>12<td>US-10-298-638-22</td><td>Sequence 22, Appl</td></td></td>	299 <td>12<td>US-10-298-638-22</td><td>Sequence 22, Appl</td></td>	12 <td>US-10-298-638-22</td> <td>Sequence 22, Appl</td>	US-10-298-638-22	Sequence 22, Appl
41 <td>41<td>31.8<td>362<td>10<td>US-09-764-864-896</td><td>Sequence 896, App</td></td></td></td></td>	41 <td>31.8<td>362<td>10<td>US-09-764-864-896</td><td>Sequence 896, App</td></td></td></td>	31.8 <td>362<td>10<td>US-09-764-864-896</td><td>Sequence 896, App</td></td></td>	362 <td>10<td>US-09-764-864-896</td><td>Sequence 896, App</td></td>	10 <td>US-09-764-864-896</td> <td>Sequence 896, App</td>	US-09-764-864-896	Sequence 896, App
42 <td>41<td>31.8<td>366<td>10<td>US-09-870-759-72</td><td>Sequence 72, Appl</td></td></td></td></td>	41 <td>31.8<td>366<td>10<td>US-09-870-759-72</td><td>Sequence 72, Appl</td></td></td></td>	31.8 <td>366<td>10<td>US-09-870-759-72</td><td>Sequence 72, Appl</td></td></td>	366 <td>10<td>US-09-870-759-72</td><td>Sequence 72, Appl</td></td>	10 <td>US-09-870-759-72</td> <td>Sequence 72, Appl</td>	US-09-870-759-72	Sequence 72, Appl
43 <td>41<td>31.8<td>366<td>12<td>US-09-751-708A-72</td><td>Sequence 72, Appl</td></td></td></td></td>	41 <td>31.8<td>366<td>12<td>US-09-751-708A-72</td><td>Sequence 72, Appl</td></td></td></td>	31.8 <td>366<td>12<td>US-09-751-708A-72</td><td>Sequence 72, Appl</td></td></td>	366 <td>12<td>US-09-751-708A-72</td><td>Sequence 72, Appl</td></td>	12 <td>US-09-751-708A-72</td> <td>Sequence 72, Appl</td>	US-09-751-708A-72	Sequence 72, Appl
44 <td>41<td>31.8<td>366<td>15<td>US-10-281-478-10</td><td>Sequence 10, Appl</td></td></td></td></td>	41 <td>31.8<td>366<td>15<td>US-10-281-478-10</td><td>Sequence 10, Appl</td></td></td></td>	31.8 <td>366<td>15<td>US-10-281-478-10</td><td>Sequence 10, Appl</td></td></td>	366 <td>15<td>US-10-281-478-10</td><td>Sequence 10, Appl</td></td>	15 <td>US-10-281-478-10</td> <td>Sequence 10, Appl</td>	US-10-281-478-10	Sequence 10, Appl
45 <td>41<td>31.8<td>414<td>12<td>US-10-032-585-7663</td><td>Sequence 7663, App</td></td></td></td></td>	41 <td>31.8<td>414<td>12<td>US-10-032-585-7663</td><td>Sequence 7663, App</td></td></td></td>	31.8 <td>414<td>12<td>US-10-032-585-7663</td><td>Sequence 7663, App</td></td></td>	414 <td>12<td>US-10-032-585-7663</td><td>Sequence 7663, App</td></td>	12 <td>US-10-032-585-7663</td> <td>Sequence 7663, App</td>	US-10-032-585-7663	Sequence 7663, App

ALIGNMENTS

RESULT 1
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PP1955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 38.8 kDa
; US-10-272-459-45

Query Match 100.0%; Score 129; DB 15; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRLGLSGVOEIKRQ 25
|||||
Db 328 WLNPKKINLADRLGLSGVOEIKRQ 352

RESULT 2
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```

; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41

Query Match          100.0%; Score 129; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 1,8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEROF
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPER,23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match          100.0%; Score 129; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4,6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPER,23AUS1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4,6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4,6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4, 6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 7
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN M
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135, 988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947, 338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397, 232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRF
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4, 6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKQ 25
956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 8
US-10-132-382-6
Sequence 6, Application US/10132382
Publication No. US20030045699A1
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132, 382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 1987
TYPE: PRF
ORGANISM: Homo sapiens
US-10-132-382-6

Query Match 37.6%; Score 48.5; DB 15; Length 1987;
Best Local Similarity 50.0%; Pred. No. 79;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24
DB 275 INLRKINLSDNHLGELPGVOSSDE 298

RESULT 9
US-10-132-382-2
Sequence 2, Application US/10132382
Publication No. US20030045699A1
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132, 382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2013
TYPE: PRF
ORGANISM: Homo sapiens
US-10-132-382-2

Query Match 37.6%; Score 48.5; DB 15; Length 2013;
Best Local Similarity 50.0%; Pred. No. 80;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24
DB 301 INLRKINLSDNHLGELPGVOSSDE 324

RESULT 10
US-10-132-382-8
Sequence 8, Application US/10132382
Publication No. US20030045699A1
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132, 382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 2014
TYPE: PRF
ORGANISM: Homo sapiens
US-10-132-382-8

Query Match 37.6%; Score 48.5; DB 15; Length 2014;
Best Local Similarity 50.0%; Pred. No. 80;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24
DB 302 INLRKINLSDNHLGELPGVOSSDE 325

RESULT 11
US-10-132-382-4
Sequence 4, Application US/10132382
Publication No. US20030045699A1
GENERAL INFORMATION:
APPLICANT: WEISS, BERTRAM
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS
FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132, 382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4

LENGTH: 2040
TYPE: PRT
ORGANISM: Homo sapiens
US-10-132-382-4

Query Match 37.6%; Score 48.5; DB 15; Length 2040;
Best Local Similarity 50.0%; Pred. No. 81;
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPKNLADRMGLSGVOEIKE 24
DB 328 INLRKLNLSDNHGLPEGVSSDE 351

RESULT 12
US-09-510-332-67
Sequence 67, Application US/09510332
Publication No. US20030022278A1
GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliott
APPLICANT: Ryba, Nick
APPLICANT: Mueller, Ken
APPLICANT: Hoon, Mark
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: T2R, a No. US20030022278A1 Family of Taste Receptors
FILE REFERENCE: 02307E-098010US
CURRENT APPLICATION NUMBER: US/09/510,332
CURRENT FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 09/393,634
PRIOR FILING DATE: 1999-09-10
NUMBER OF SEQ ID NOS: 172
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 67
LENGTH: 309
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human T2R41 (hGR41)
NAME/KEY: MOD_RES
LOCATION: (253)
OTHER INFORMATION: Xaa - any amino acid
US-09-510-332-67

Query Match 34.9%; Score 45; DB 11; Length 309;
Best Local Similarity 36.4%; Pred. No. 34;
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKE 22
35 WVNTRKISSADQIILTAIAVSRY 56

RESULT 13
US-09-764-864-1535
Sequence 1535, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1535
LENGTH: 325
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (309)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (322)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1535

Query Match 34.9%; Score 45; DB 10; Length 325;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPKKINLADRMGLSGVOEIKE 24
DB 210 NPKRLHLTGMLVLLSEEEIQO 231

RESULT 14
US-09-938-803-16
Sequence 16, Application US/09938803
Patent No. US20020076762A1
GENERAL INFORMATION:
APPLICANT: Yue, Henry
APPLICANT: Tang, Y. Tom
APPLICANT: Lal, Preeti
APPLICANT: Reddy, Roopa
APPLICANT: Baughn, Mariah R.
APPLICANT: Yang, Junning
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS
FILE REFERENCE: PF-0695 US
CURRENT APPLICATION NUMBER: US/09/938,803
CURRENT FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 09/311,894
PRIOR FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 40
SOFTWARE: PERL Program
SEQ ID NO 16
LENGTH: 357
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte Clone 2805526
US-09-938-803-16

Query Match 34.9%; Score 45; DB 9; Length 357;
Best Local Similarity 45.5%; Pred. No. 41;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPKKINLADRMGLSGVOEIKE 24
DB 174 NPKRLHLTGMLVLLSEEEIQO 195

RESULT 15
US-09-764-864-1118
Sequence 1118, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1118
LENGTH: 371
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (309)
Query Match 34.9%; Score 45; DB 10; Length 371;

Best Local Similarity 45.5%; Pred. No. 43;
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
OY 3 NPKINLADRMGLSGVOEIKE 24
||||:| ||| :||:
Db 188 NPKLHLRTIGMLVLLSEETIOO 209

Search completed: October 1, 2003, 10:37:54
Job time : 30.1471 secs

THIS PAGE BLANK (USPTO)